

ADA080350

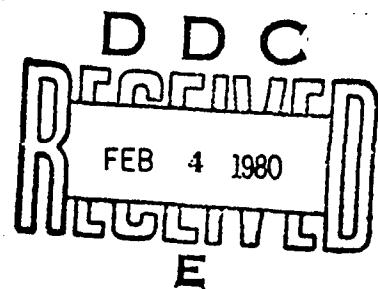
LEVEL

RESEARCH REPORT 79-4



A NEW FAMILY OF MODELS FOR THE  
MULTIPLE-CHOICE ITEM

FUMIKO SAMEJIMA



DEPARTMENT OF PSYCHOLOGY  
UNIVERSITY OF TENNESSEE  
KNOXVILLE, TENN. 37916

DECEMBER, 1979

DDC FILE COPY

Prepared under the contract number NOOO14-77-C-360,  
NR 150-402 with the  
Personnel and Training Research Programs  
Psychological Sciences Division  
Office of Naval Research

Approved for public release; distribution unlimited.  
Reproduction in whole or in part is permitted for  
any purpose of the United States Government.

80 2 1 047

## Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE			READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Research Report 79-4	2. GOVT ACCESSION NO. <i>(14)RR-79-4</i>	3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) A New Family of Models for the Multiple-Choice Item	5. TYPE OF REPORT & PERIOD COVERED Technical Report		
6. PERFORMING ORG. REPORT NUMBER			
7. AUTHOR(s) Dr. Fumiko Samejima	8. CONTRACT OR GRANT NUMBER(s) <i>(15)N00014-77-C-0360</i>		
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology University of Tennessee Knoxville, Tennessee 37916	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE: 61153N; PROJ: RR 042-04 TA: RR 042-04-01 WU: NR 150-402		
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, VA 22217	12. REPORT DATE <i>(11)19 December 79</i>		
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) <i>(16)RR04204</i> <i>(17)RR0420401</i>	13. NUMBER OF PAGES 140 <i>(12)142</i>		
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States government.	15. SECURITY CLASS. (of this Report) Unclassified		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE		
18. SUPPLEMENTARY NOTES			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Operating Characteristic Estimation Tailored Testing Latent Trait Theory			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  (Please see reverse side)			

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

  
The three-parameter logistic model has been used by many researchers as the model for the multiple-choice item, regardless of the fact that for most multiple-choice test items the examinee's behavior does not follow the knowledge or random guessing principle, upon which the model is based.

Estimation of the operating characteristics without assuming any mathematical forms has been pursued by the present author and many combinations of methods and approaches have been produced. The application of these methods for many empirical data will enable us to discover the operating characteristics of multiple-choice items, and, eventually, lead us to more meaningful models than the three-parameter logistic model. While the research in this direction is in process, however, it will be helpful if some other model or models, which is based upon a sounder rationale than the knowledge or random guessing principle, is proposed. 

In the present paper, a family of models for the multiple-choice item is proposed for this purpose. These models are built in consideration of the behavior of distractors of the multiple-choice item, as well as the examinee's random guessing. One incentive for proposing these models is Shiba's research which includes the construction of a vocabulary test, and which I came across while I was doing research in Tokyo, Japan in summer, 1979.

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

## A NEW FAMILY OF MODELS FOR THE MULTIPLE-CHOICE ITEM

### ABSTRACT

The three-parameter logistic model has been used by many researchers as the model for the multiple-choice item, regardless of the fact that for most multiple-choice test items the examinee's behavior does not follow the knowledge or random guessing principle, upon which the model is based.

Estimation of the operating characteristics without assuming any mathematical forms has been pursued by the present author and many combinations of methods and approaches have been produced. The application of these methods for many empirical data will enable us to discover the operating characteristics of multiple-choice items, and, eventually, lead us to more meaningful models than the three-parameter logistic model. While the research in this direction is in process, however, it will be helpful if some other model or models, which is based upon a sounder rationale than the knowledge or random guessing principle, is proposed.

In the present paper, a family of models for the multiple-choice item is proposed for this purpose. These models are built in consideration of the behavior of distractors of the multiple-choice item, as well as the examinee's random guessing. One incentive for proposing these models is Shiba's research which includes the construction of a vocabulary test, and which I came across while I was doing research in Tokyo, Japan in summer, 1979.

---

The research was conducted at the principal investigator's laboratory, 409 Austin Peay Hall, Department of Psychology, University of Tennessee, Knoxville, Tennessee. Those who worked for her as research assistants include Paul S. Changas and Philip S. Livingston. Typing and data organization were helped by Nancy Jayne Taylor, Deusdedit Furlan and Tamra Gordon.

## TABLE OF CONTENTS

	Page
I Introduction	1
II Normal Ogive Model on the Graded Response Level and Bock's Multinomial Model	5
III A New Family of Models for the Multiple-Choice Test Item	14
IV Basic Functions and Information Functions of the Multiple-Choice Item	62
V Qualities that Distinguish Good Test Items from Bad Ones	90
VI Discussion and Conclusions	96
References	99
Appendix I	103
Appendix II	119

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DDC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	<input type="checkbox"/>
By _____	
Distribution/	
Availability Codes	
Dist	Available/or special
A	

## I Introduction

The three-parameter normal ogive, or logistic, model (Birnbaum, 1968) has been widely used for the multiple-choice test item among psychometricians and other researchers in mental measurement (e.g., Bejar, I. I., D. J. Weiss and G. G. Kingsbury, 1977, Hambleton, R. K. and J. Gifford, 1979, McBride, J. R., 1977, Reckase, M. D., 1977, Swaminathan, H. and J. Gifford, 1979, Sympson, J. B., 1977, Urry, V. W., 1977, Warm, T. A., 1978). The model is based upon the knowledge or random guessing principle, i.e., the examinee either knows the answer, or guesses randomly among the alternatives. It is alarming to note, however, that, in spite of its unusual popularity, none of the researchers have even tried to validate the model, but adopted it rather blindly, except for Lord (Lord, 1970). Experienced test constructors try to include wrong, but plausible, answers among the alternatives of multiple-choice items, which are called distractors, so as not to make the correct answers too conspicuous and destroy the quality of questions. It should be noted that we need some higher mental processes other than random guessing to recognize the plausibility of a distractor, and to be attracted to it. It is contradictory, therefore, to apply the three-parameter normal ogive, or logistic, model for multiple-choice items with distractors, although many researchers do.

Bock has developed a multinomial model (Bock, 1972) for the multiple-choice item taking a completely different standpoint. He postulated his rationale assuming a response tendency for each

alternative, including the correct answer, and assumed a normal distribution for the conditional distribution of each response tendency, given ability, and the invariance of the ratio of the conditional probabilities with which two alternatives are selected in preference to each other, regardless of the set of alternatives they are placed in. In Bock's model, no considerations are given for the examinee's random guessing behavior. It is assumed that every examinee seriously compares each alternative with each other, and selects his answer.

The homogeneous and heterogeneous cases of the graded response model have been proposed by Samejima (Samejima, 1972), and later the model has further been expanded to the continuous response model (Samejima, 1973a). Typical examples of models in the homogeneous case on the graded response level are the normal ogive and the logistic models, which were originally proposed as models for binary items (cf. Lord and Novick, 1968). This family of models was built and proposed, mainly, for the purpose of extracting a greater amount of information from free-response test items. Many researchers have thought that they are for the multiple-choice item, however. If we take this latter standpoint, then we shall have to say that these models are close to the multinomial model in the sense that the idea of random guessing is completely missing.

The role of the wrong answers as alternatives in the multiple-choice item is simply to provide noise, as it is viewed by the three-parameter normal ogive, or logistic, model. The multiple-choice item is, therefore, nothing but a "blurred image" of the free-response

item. On the other hand, in Bock's model or in the normal ogive, or logistic, model on the graded response level, the multiple-choice test item is much more than a poor substitute for the free-response item, i.e., the item which provides us with a greater amount of information and a higher accuracy of ability estimation, than the binary, free-response item.

When we consider the fact that most well-constructed multiple-choice items have distractors, the multinomial model or one of the graded response models looks more reasonable to adopt than the three-parameter normal ogive, or logistic, model. It should be noted, however, that we cannot completely ignore the examinee's random guessing behavior, since most examinees depend upon random guessing as the last resort, when a strong pressure for success is present, as is the case in many testing situations. We must conclude, therefore, that neither the three-parameter normal ogive, or logistic, model nor the multinomial or graded response model serves our purpose of interpreting the multiple-choice item properly, when it is given in a typical testing session.

Estimation of the operating characteristics of the graded response categories has been investigated by Samejima (Samejima, 1977b, 1977c, 1978a, 1978b, 1978c, 1978d, 1978e, 1978f), without assuming any mathematical forms. One of the incentives when the author started this part of research in latent trait theory is to investigate the operating characteristics of the distractors in the multiple-choice item, using the estimation methods thus developed. When this is realized, we can make use of the information given by

the distractors, as well as the one given by the correct answer, and create a more efficient ability estimation by means of multiple-choice items. Under the circumstances, this approach seems to be the most productive, and scientific, one. It will be beneficial, however, if we can propose some other family of models that has more reasonable rationale for the multiple-choice test item, to satisfy the immediate need of replacing the three-parameter normal ogive, or logistic, model.

For this purpose, in the present paper, a new family of models for the multiple-choice item is formulated and proposed. These models are a natural consequence of combining the higher mental process of recognizing the plausibility of distractors and random guessing. In addition to this proposal, desirable qualities of multiple-choice test items are discussed in the light of the new family of models.

## II Normal Ogive Model on the Graded Response Level and Bock's Multinomial Model

The normal ogive model, which was originally introduced for a binary, free-response item, has been expanded to fit a more general case, in which the item is graded into more than two item score categories (Samejima, 1969, 1972). In this chapter, we shall compare this model with Bock's multinomial model.

It has been pointed out (Samejima, 1972) that, although Bock's model was originally developed for nominal categories, i.e., the categories which are not ordered among themselves, it can be considered as a model in the heterogeneous case of the graded response level.

Let  $\theta$  denote ability, or latent trait, which is assumed to be uni-dimensional, such that

$$(2.1) \quad -\infty < \theta < \infty .$$

Let  $g$  be a multiple-choice item,  $h$ ,  $i$  or  $k$  be one of its  $m_g$  alternatives, and  $x_{hg}$ ,  $x_{ig}$  and  $x_{kg}$  be the response tendency for the alternative,  $h$ ,  $i$  or  $k$ . When any two alternatives,  $h$  and  $k$ , are compared alone, the probability with which  $h$  is chosen in preference to  $k$  is assumed to be a function of ability  $\theta$ , and is denoted by  $\pi_{hk}(\theta;g)$ . Thus we can write

$$(2.2) \quad \pi_{hk}(\theta;g) + \pi_{kh}(\theta;g) = 1 .$$

When the comparison is made among  $m_g$  ( $\geq 2$ ) alternatives, the conditional probability with which the alternative  $h$  is chosen

in preference to all the other  $(m-1)$  alternatives, given  $\theta$ , is denoted by  $P_h(\theta;g)$ , and we have

$$(2.3) \quad \sum_{h=1}^m P_h(\theta;g) = 1.$$

We shall define a variable  $X_{hk;g}$ , such that

$$(2.4) \quad X_{hk;g} = X_{hg} - X_{kg},$$

i.e., the difference between the two response tendencies,  $X_{hg}$  and  $X_{kg}$ .

Hereafter, for simplicity, we shall drop the subscript  $g$ , whenever it is clear that we are dealing with only one multiple-choice item. Thus, in such a case,  $\pi_{hk}(\theta)$  is used for  $\pi_{hk}(\theta;g)$ ,  $X_{hk}$  for  $X_{hk;g}$ , and so forth.

In the multinomial model, it is assumed that: 1) the conditional distribution of  $X_k$ , given  $\theta$ , is normal, with  $\mu_k(\theta;g)$  or  $\mu_k(\theta)$ , and  $\sigma_k(\theta;g)$ , or  $\sigma_k(\theta)$ , as the two parameters; 2)  $X_k$ 's are conditionally, mutually independent, given  $\theta$ ; and 3) the ratio of the probabilities with which the two alternatives are chosen, respectively, is invariant for the set of alternatives among which the two alternatives are compared. Thus for the third assumption we can write

$$(2.5) \quad P_h(\theta)/P_k(\theta) = \pi_{hk}(\theta)/\pi_{kh}(\theta).$$

From the first two of these assumptions, it is derived that the conditional distribution of  $X_{hk}$ , given  $\theta$ , is also normal,

with  $\mu_{hk}(\theta;g)$ , or  $\mu_{hk}(\theta)$ , and  $\sigma_{hk}(\theta;g)$ , or  $\sigma_{hk}(\theta)$ , as the two parameters, which are given by

$$(2.6) \quad \mu_{hk}(\theta) = \mu_h(\theta) - \mu_k(\theta),$$

and

$$(2.7) \quad \sigma_{hk}(\theta) = [\sigma_h^2(\theta) + \sigma_k^2(\theta)]^{1/2}.$$

We can also write for  $\pi_{hk}(\theta)$  and  $\pi_{kh}(\theta)$  such that

$$(2.8) \quad \pi_{hk}(\theta) = (2\pi)^{-1/2} \sigma_{hk}(\theta)^{-1} \int_0^\infty \exp[-\{x_{hk} - \mu_{hk}(\theta)\}^2 / \{2\sigma_{hk}^2(\theta)\}] dx_{hk}$$

and

$$(2.9) \quad \pi_{kh}(\theta) = (2\pi)^{-1/2} \sigma_{hk}(\theta)^{-1} \int_{-\infty}^0 \exp[-\{x_{hk} - \mu_{hk}(\theta)\}^2 / \{2\sigma_{hk}^2(\theta)\}] dx_{hk}.$$

Now we shall use the logistic approximation to the normal distribution function, which is, with  $D = 1.7$ , given by

$$(2.10) \quad (2\pi)^{-1/2} \int_{-\infty}^u e^{-v^2/2} dv \doteq [1 + \exp\{-Du\}]^{-1}.$$

Thus we obtain from (2.6), (2.7), (2.8), (2.9) and (2.10)

$$\begin{aligned} (2.11) \quad \pi_{hk}(\theta) / \pi_{kh}(\theta) &\doteq [1 + \exp\{D\mu_{hk}(\theta) / \sigma_{hk}(\theta)\}] \\ &\quad [1 - \{1 + \exp\{D\mu_{hk}(\theta) / \sigma_{hk}(\theta)\}\}]^{-1} \\ &= \exp[D\mu_{hk}(\theta) / \sigma_{hk}(\theta)] \\ &= \exp[\mu_h(\theta) - \mu_k(\theta) / \{\sigma_h^2(\theta) + \sigma_k^2(\theta)\}^{1/2}]. \end{aligned}$$

From (2.5) and (2.11) we can write

$$(2.12) \quad [P_k(\theta)]^{-1} = \sum_{i=1}^m [P_i(\theta) / P_k(\theta)] = \sum_{i=1}^m [\pi_{ik}(\theta) / \pi_{ki}(\theta)]$$

$$\doteq \sum_{i=1}^m \exp[D\{\mu_i(\theta) - \mu_k(\theta)\}/\{\sigma_i^2(\theta) + \sigma_k^2(\theta)\}^{1/2}] ,$$

and then

$$(2.13) \quad P_h(\theta) = P_k(\theta) [\pi_{hk}(\theta)/\pi_{kh}(\theta)] \\ \doteq \exp[D\{\mu_h(\theta) - \mu_k(\theta)\}/\{\sigma_h^2(\theta) + \sigma_k^2(\theta)\}^{1/2}] \\ [ \sum_{i=1}^m \exp[D\{\mu_i(\theta) - \mu_k(\theta)\}/\{\sigma_i^2(\theta) + \sigma_k^2(\theta)\}^{1/2}] ]^{-1} ,$$

for  $h = 1, 2, \dots, m$ . Note that  $k$  is an arbitrarily chosen, fixed alternative.

If we add two other assumptions such that: 4) the regression of the response tendency  $X_h$  ( $h=1, 2, \dots, m$ ) is linear; and 5) the conditional variance of  $X_h$ , given  $\theta$ , is constant, i.e.,

$$(2.14) \quad \mu_h(\theta) = a_h^* \theta + c_h^*$$

and

$$(2.15) \quad \sigma_h^2(\theta) = \sigma_h^2 ,$$

then we can write

$$(2.16) \quad D[\mu_h(\theta) - \mu_k(\theta)]/\{\sigma_h^2(\theta) + \sigma_k^2(\theta)\}^{1/2} = a_h \theta + c_h ,$$

where

$$(2.17) \quad a_h = D(a_h^* - a_k^*)/\{\sigma_h^2 + \sigma_k^2\}^{1/2}$$

and

$$(2.18) \quad c_h = D(c_h^* - c_k^*)/\{\sigma_h^2 + \sigma_k^2\}^{1/2} .$$

Substituting (2.16) into (2.13), we obtain

$$(2.19) \quad P_h(\theta) = \exp[a_h\theta + c_h] \left[ \sum_{i=1}^m \exp[a_i\theta + c_i] \right]^{-1}.$$

Thus (2.19) specifies the operating characteristic of the category  $h$  in the multinomial model. Note that both  $a_i$ 's and  $c_i$ 's in (2.19) are of arbitrary origins, for we have for arbitrary  $d$  and  $e$

$$(2.20) \quad P_h(\theta) = \exp[a_h\theta + c_h] \exp[d\theta + e] \left[ \sum_{i=1}^m \exp[d\theta + e] \exp[a_i\theta + c_i] \right]^{-1}$$

$$= \exp[(a_h+d)\theta + (c_h+e)] \left[ \sum_{i=1}^m \exp[(a_i+d)\theta + (c_i+e)] \right]^{-1}.$$

While in the multinomial model we assume  $m$  different response tendencies and their conditional independence, and the invariance of the ratio of the two probabilities of alternative selection, in the normal ogive model on the graded response level, we assume that there exists a single response tendency, or item variable,  $X_g$ , or  $X$ , behind the selection of any one of the  $m$  alternatives, and the conditional distribution of  $X$ , given  $\theta$ , is normal, with  $\mu(\theta)$  and  $\sigma(\theta)$  as the two parameters. In addition to this first assumption, we also assume for the normal ogive model that: 2) the whole dimension of the item variable  $X$  is divided into  $m$  subintervals; and 3) the alternative  $h$  will be selected if the examinee's response tendency is in the subinterval assigned to that category. We can write

$$(2.21) \quad P_h(\theta) = [2\pi]^{-1/2} [\sigma(\theta)]^{-1} \int_{Y_{h-1}}^{Y_h} \exp[-(u-\mu(\theta))^2 / \{2\sigma(\theta)^2\}] du,$$

where  $\gamma_h$  is the upper endpoint of the subinterval of  $X$  which is assigned to the category  $h$ , and we have

$$(2.22) \quad \begin{cases} \gamma_0 = -\infty \\ \gamma_m = \infty \end{cases}$$

The additional two assumptions, 4) and 5), for the multinomial model, which are formulated by (2.14) and (2.15), respectively, are also adopted for the item variable  $X$  in the normal ogive model. Thus we can write for the conditional expectation, or regression, of  $X$  on  $\theta$  and the conditional variance of  $X$ , given  $\theta$ ,

$$(2.23) \quad \mu(\theta) = a^*\theta + c^* ,$$

and

$$(2.24) \quad \sigma^2(\theta) = \sigma^2 .$$

Substituting (2.23) and (2.24) into (2.21), we obtain

$$(2.25) \quad \begin{aligned} p_h(\theta) &= [2\pi]^{-1/2} \sigma^{-1} \int_{\gamma_{h-1}}^{\gamma_h} \exp[-(u-a^*\theta-c^*)^2/(2\sigma^2)] du \\ &= [2\pi]^{-1/2} \int_{(\gamma_{h-1}-a^*\theta-c^*)/\sigma}^{(\gamma_h-a^*\theta-c^*)/\sigma} \exp[-t^2/2] dt \\ &= [2\pi]^{-1/2} \int_{(a^*\theta+c^*-\gamma_{h-1})/\sigma}^{(a^*\theta+c^*-\gamma_h)/\sigma} \exp[-t^2/2] dt , \end{aligned}$$

where

$$(2.26) \quad t = (u-a^*\theta-c^*)/\sigma .$$

We define the item parameters,  $a_g$ , or  $a$ , and  $b_{hg}$ , or  $b_h$ , such that

$$(2.27) \quad a = a^*/\sigma$$

and

$$(2.28) \quad b_h = (\gamma_h - c^*)/a^* ,$$

where

$$(2.29) \quad \begin{cases} b_0 = -\infty \\ b_m = \infty \end{cases} .$$

Substituting (2.27) and (2.28) into (2.25), we obtain for the normal ogive model on the graded response level,

$$(2.30) \quad P_h(\theta) = [2\pi]^{-1/2} \int_{a(\theta-b_h)}^{a(\theta-b_{h-1})} \exp[-t^2/2] dt .$$

We have seen in the preceding paragraphs that in both the normal ogive model on the graded response level and Bock's multinomial model the normal assumption is made for the conditional distribution of the response tendency, given ability  $\theta$ . The biggest difference between the two models is that, in the normal ogive model, a single item variable is assumed behind the examinee's selection behavior, whereas, in the multinomial model, a separate response tendency is assumed for each of the  $m$  alternatives. The decision as to which model should be adopted should depend upon the psychological reality behind our data.

It should be pointed out that in both models the operating characteristic is strictly decreasing in  $\theta$  for the category 1, unimodal for the categories, 2 through  $(m-1)$ , and strictly increasing in  $\theta$  for the category  $m$ , and the modal points of these operating characteristics are in the ascending order, provided that

$$(2.31) \quad a_1 < a_2 < \dots < a_m$$

in the multinomial model (cf. Samejima, 1972). It should also be noted that in neither model is the effect of random guessing accounted for. The application of these models for the multiple-choice test item must be restricted, therefore, to the case where the supervision of examinees is well conducted and guessing is strongly discouraged in the instructions.

Another interesting difference between the two models is that the normal ogive model requires ordering the response categories a priori, while the multinomial model does not, in estimating the item parameters. This ordering is a fairly easy process, however. One of the good methods of ordering the alternatives of the multiple-choice item when we use the normal ogive model may be the following.

- (1) Treat all the items as binary items, and estimate each examinee's ability by the maximum likelihood estimation.

This process is facilitated if we use the logistic approximation (cf. Birnbaum, 1968), whose item characteristic function,

$P_g(\theta)$ , or  $P(\theta)$ , is given by

$$(2.32) \quad P(\theta) = [1+\exp\{-Da(\theta-b_{m-1})\}]^{-1}.$$

(2) Find out the sample mean of the maximum likelihood estimate of ability for each subgroup of examinees, who have chosen a specific alternative.

(3) Order the alternatives according to the sample means obtained in (2).

We can expand the logistic approximation to the normal ogive model further for the category  $h$ , and define the logistic model such that, for  $h = 1, 2, \dots, m$ ,

$$(2.33) \quad P_h(\theta) = \frac{[1-\exp\{-Da(b_h - b_{h-1})\}][1+\exp\{-Da(\theta - b_{h-1})\}]^{-1}}{[1+\exp\{Da(\theta - b_h)\}]^{-1}},$$

where  $b_0$  and  $b_m$  satisfy (2.29). This model has similar characteristics as the normal ogive model, although it also has interesting differences (cf. Samejima, 1972).

The three models described in the previous paragraphs of this chapter can be applied for the multiple-choice item, but only in a restricted way. As we have seen, the examinee's guessing behavior is not considered in the rationale behind these models. Furthermore, it is unrealistic to assume a strictly decreasing function for the operating characteristic of one of the wrong answers. This problem will be solved by modifying the models, which will be presented in the following chapter.

### III A New Family of Models for the Multiple-Choice Test Item

Suppose that our multiple-choice test item is constructed well enough to provide us with  $(m-1)$  distractors, which have some plausibility to attract examinees as correct answers. Suppose, further, that there is some simple statistical relationship between each distractor and ability  $\theta$ , i.e., the conditional probability with which the examinee chooses the distractor  $h$  as the correct answer in comparison with all the other  $(m-1)$  alternatives, given  $\theta$ , which increases in  $\theta$  up to a certain level of  $\theta$ , and then decreases in  $\theta$ . This implies that there possibly are individuals who are not even good enough to recognize the plausibility of any distractor as a correct answer. Suppose that the conditional probability with which the examinee does neither solve the problem nor recognize the plausibility of any distractor, given  $\theta$ , is strictly decreasing in ability  $\theta$ . If the item characteristic function, which is the conditional probability, given  $\theta$ , with which the examinee selects the correct answer, is strictly increasing in ability  $\theta$  with zero and unity as its two asymptotes, and if, in addition, the two asymptotes of the "plausibility" function for each of the  $(m-1)$  distractors are uniformly zero, and those of the conditional probability for the "no recognition" category are unity and zero, respectively, then we will notice that the type of models which includes both the normal ogive model on the graded response level and the multinomial model may be appropriate for our test item.

In situations like the one described in the preceding paragraph, the type of models is suitable only if the supervision

is strict and the examinees are extremely discouraged to guess when they do not know the right answer. It may be more realistic to assume, however, that in most testing situations the pressure for success is so strong that the examinees do guess when they have no idea about the correct answer. Suppose that these examinees guess randomly, and select one of the  $m_g$  alternatives with equal probability. Thus we obtain a new family of models, which includes modified forms of such models as the normal ogive and logistic models on the graded response level and the multinomial model.

Let  $P_{x_g}(\theta)$  be the operating characteristic of the graded response category  $x_g (=0,1,2,\dots,m_g)$ , whose mathematical form is given as  $P_h(\theta;g)$  in (2.19), (2.30) or (2.33), or of any other model of similar characteristics. For convenience, we shall call these models as models of Type I on the graded response level, just as we did on the dichotomous response level (Samejima, 1979). To be specific, models of Type I are those which satisfy the following.

- (1)  $P_{x_g}(\theta)$  is strictly decreasing in  $\theta$ , with unity and zero as its two asymptotes, for  $x_g = 0$ .
- (2)  $P_{x_g}(\theta)$  is unimodal with zero as its two asymptotes, for  $x_g = 1,2,\dots,(m_g - 1)$ .
- (3)  $P_{x_g}(\theta)$  is strictly increasing in  $\theta$ , with zero and unity as its two asymptotes, for  $x_g = m_g$ .

The above conditions for Type I models also imply that  $\sum_{s=x_g}^{m_g} P_s(\theta)$  is strictly increasing in  $\theta$  with zero and unity as its two  $m_g$  asymptotes, for  $x_g = 1,2,\dots,m_g$ .

We use this additional response category  $x_g = 0$  for those who have no idea at all of the correct answer in the multiple-choice situation. Thus the probability with which the examinee belongs to this category is strictly decreasing in  $\theta$ , with unity and zero as its two asymptotes. We assume that the  $(m_g - 1)$  distractors of the multiple-choice item  $g$  have an implicit order among themselves, and the response categories  $x_g = 1, 2, \dots, (m_g - 1)$  are used for the distractors. Thus the operating characteristics of the distractors are unimodal, with zero as their two asymptotes, respectively. The other category,  $x_g = m_g$ , is used for the correct answer, and its operating characteristic is strictly increasing in  $\theta$  with zero and unity as its two asymptotes. Since, in reality, the examinees who belong to the category,  $x_g = 0$ , are assumed to guess randomly, however, the operating characteristic for this response category disappears, and those of the other categories, or the  $m_g$  alternatives, are affected by this random guessing. The operating characteristic of the alternative  $h$  can be written, therefore, such that

$$(3.1) \quad P_h(\theta; g) = P_{x_g}(0; x_g = h) + (1/m_g)P_{x_g}(\theta; x_g = 0) .$$

Thus (3.1) specifies the new family of models for the multiple-choice item. When  $P_{x_g}(\theta)$  follows the normal ogive model on the graded response level,  $P_h(\theta; g)$ , or  $P_h(\theta)$ , takes on a form such that

$$(3.2) \quad P_h(\theta) = (2\pi)^{-1/2} \left[ \int_{a(\theta-b_{h+1})}^{a(\theta-b_h)} e^{-u^2/2} du + (1/m) \int_{a(\theta-b_1)}^{\infty} e^{-u^2/2} du \right] ,$$

where  $a > 0$ , and

$$(3.3) \quad -\infty < b_1 < b_2 < \dots < b_m < b_{m+1} = \infty .$$

For simplicity, we shall call it Model A of Type I for the multiple-choice item. When  $P_{x_g}(\theta)$  in (3.1) is specified by the logistic model on the graded response level, we can write

$$(3.4) \quad P_h(\theta) = \frac{[1-\exp\{-Da(b_{h+1}-b_h)\}][1+\exp\{-Da(\theta-b_h)\}]}{[1+\exp\{Da(\theta-b_{h+1})\}] + [m[1+\exp\{Da(\theta-b_1)\}]]^{-1}} ,$$

where  $a > 0$ , and the inequality (3.3) holds. We shall call it Model B of Type I for the multiple-choice item. When  $P_{x_g}(\theta)$  in (3.1) is given by the operating characteristic of the category  $x_g$  in the multinomial model, we obtain

$$(3.5) \quad P_h(\theta) = \frac{[\exp\{a_h\theta+c_h\} + (1/m)\exp\{a_0\theta+c_0\}][\sum_{i=0}^m \exp\{a_i\theta+c_i\}]}{\sum_{i=0}^m \exp\{a_i\theta+c_i\}}^{-1} ,$$

where

$$(3.6) \quad a_0 < a_1 < a_2 < \dots < a_m .$$

We shall call it Model C of Type I for the multiple-choice item, or Bock-Samejima model for the multiple-choice item.

For the purpose of illustration, Figure 3-1 presents the operating characteristics of the six response categories, following the normal ogive model on the graded response level, with  $a_g = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -0.50$ ,  $b_3 = 0.00$ ,  $b_4 = 0.75$  and  $b_5 = 1.25$ . The modal point of the operating characteristic of each of the  $(m-1)$  intermediate categories is given by  $(b_h + b_{h+1})/2$  (Samejima,

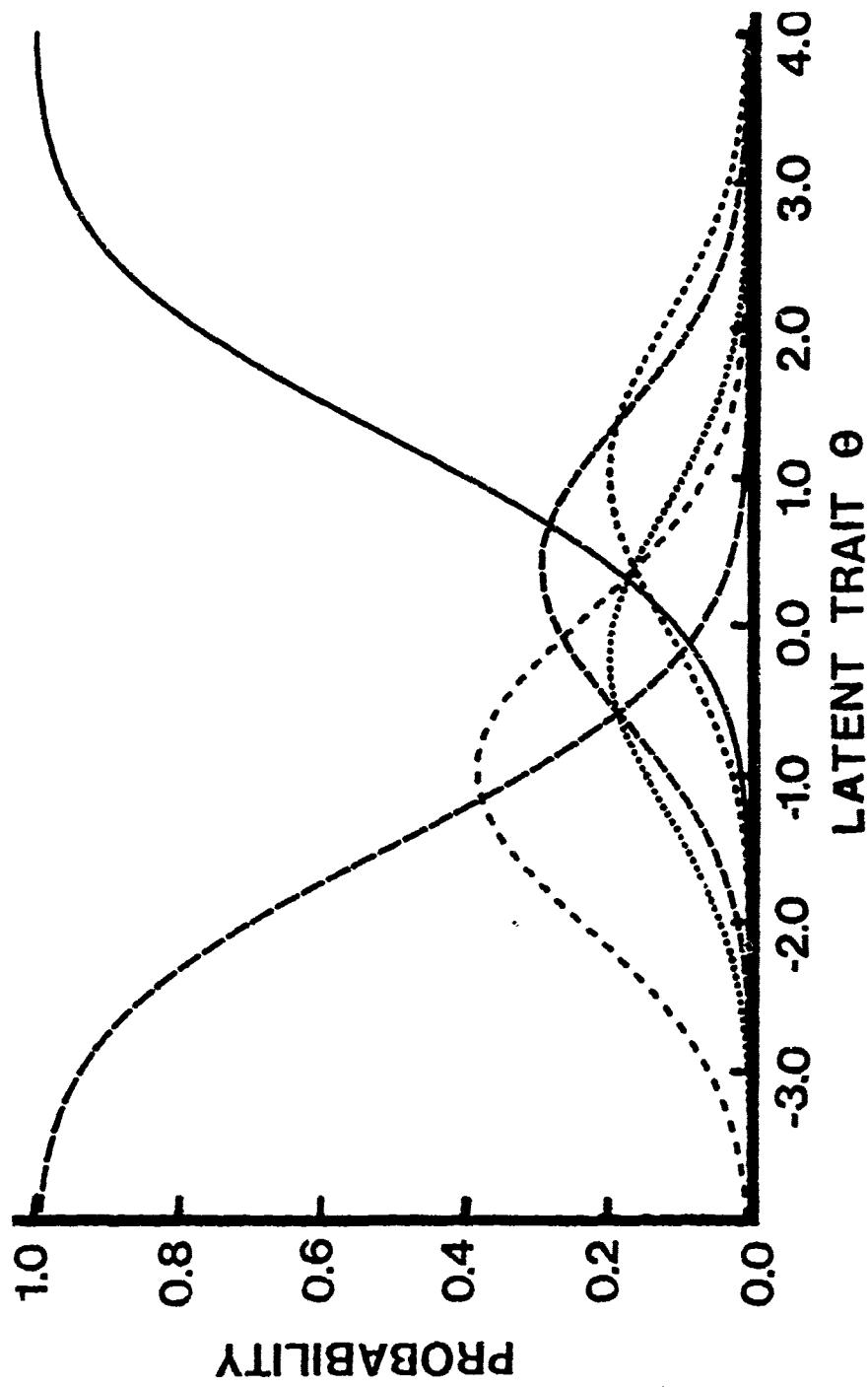


FIGURE 3-1

Operating Characteristics of Six Item Response Categories Following the Normal Ogive Model, with  $a_g = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -0.50$ ,  $b_3 = 0.00$ ,  $b_4 = 0.75$  and  $b_5 = 1.25$ .

1969), and this will help the reader to identify each operating characteristic in Figure 3-1. Figure 3-2 presents the corresponding operating characteristics of the five alternatives following Model A of Type I for the multiple-choice item. We can see that these curves are no longer symmetric for  $h = 1, 2, 3$  and  $4$ , and they are not even unimodal for  $h = 2, 3$  and  $4$ . It is indicated in the figure that the asymptotes of these operating characteristics at  $\theta = -\infty$  are uniformly  $0.2$ , or  $1/m$ . To make this contrast clear, Figure 3-5 presents the five pairs of operating characteristics in the normal ogive model and Model A, with dotted and solid lines, for the alternatives, 1 through 5, respectively. We can see in these graphs that the modal point is shifted to the negative direction in Model A, for each of the alternatives, 1 through 4, and the amount of shift is greater for lower categories. We also observe that for the correct answer, or  $h = 5$ , the operating characteristic decreases in  $\theta$  for a certain interval of  $\theta$ , and then increases in  $\theta$ . This shape is similar to the one observed for a mathematics test item (Lord and Novick, 1968, Figure 16.4.1), for which a rather crude approximation to the item characteristic function, i.e., the percentage correct plotted against the total test score, is used. It should be noted, moreover, that, if the curve is truncated around  $\theta = -0.8$ , then it looks as if the operating characteristic had a lower asymptote at  $\theta = -\infty$  than  $1/m$ . If we combine this with the fact that this operating characteristic converges to the item characteristic function of the three-parameter normal ogive model as  $b_1, b_2, \dots, b_{m-1}$  tend to  $b_m$ , then we will notice that

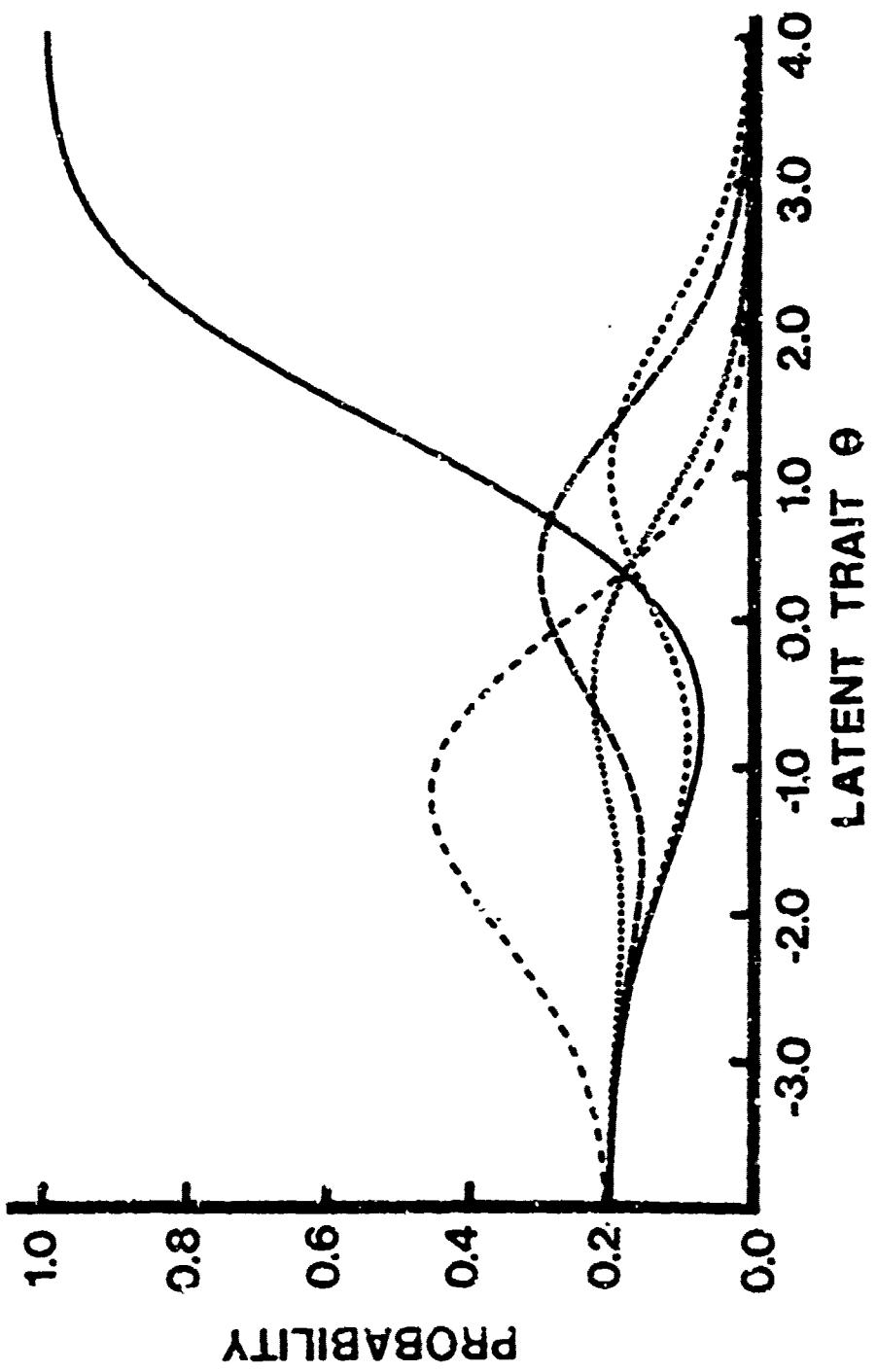


FIGURE 3-2

Operating Characteristics of Five Alternatives following the Model A of Type I for the Multiple-Choice Item. The Parameters Are:  $a = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -0.50$ ,  $b_3 = 0.00$ ,  $b_4 = 0.75$  and  $b_5 = 1.25$ .

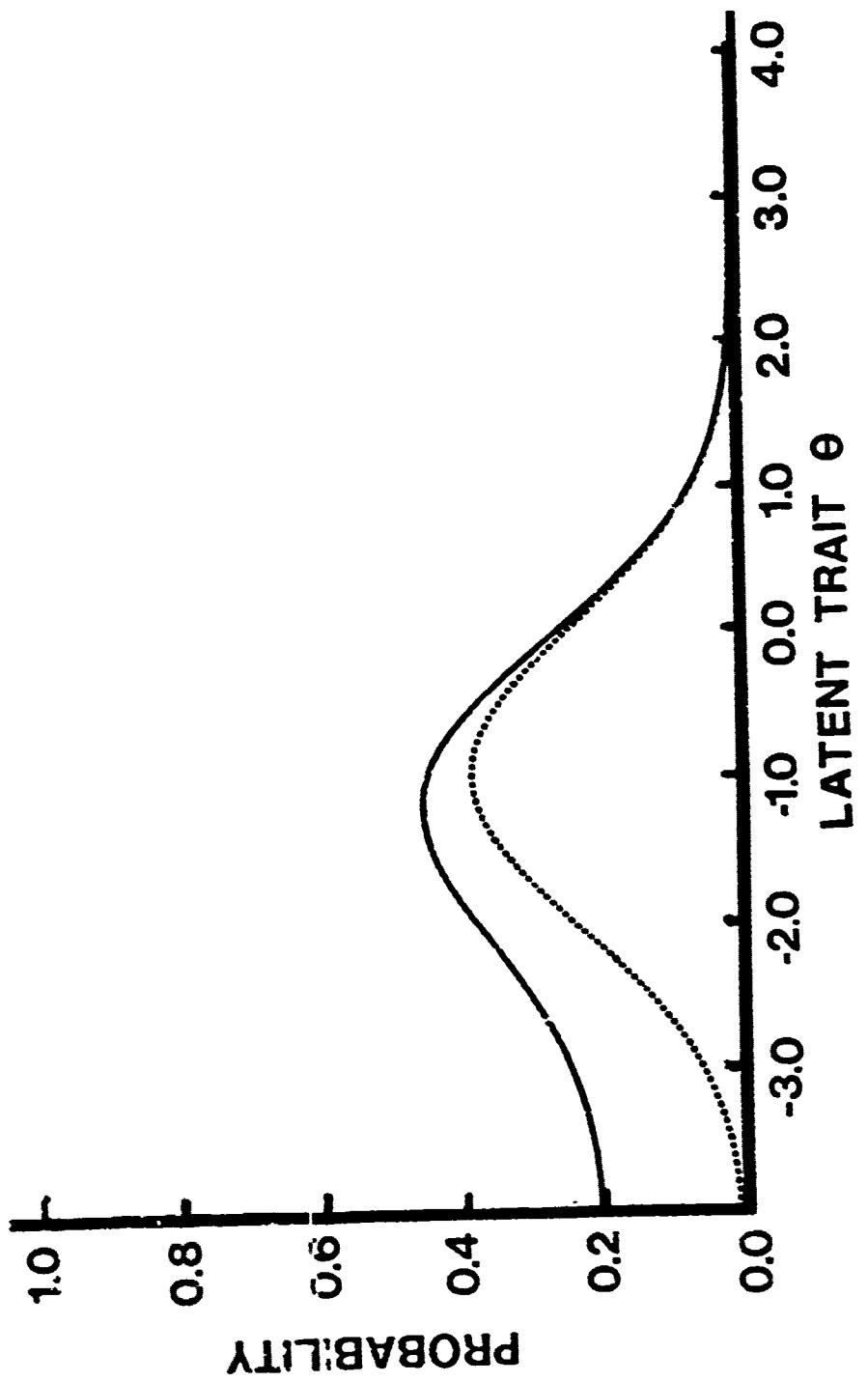


FIGURE 3-3

Comparison of the Two Operating Characteristics in the Normal Ogive Model  
(Dotted Curve) and in Model A of Type I (Solid Curve).  $h = 1$ .

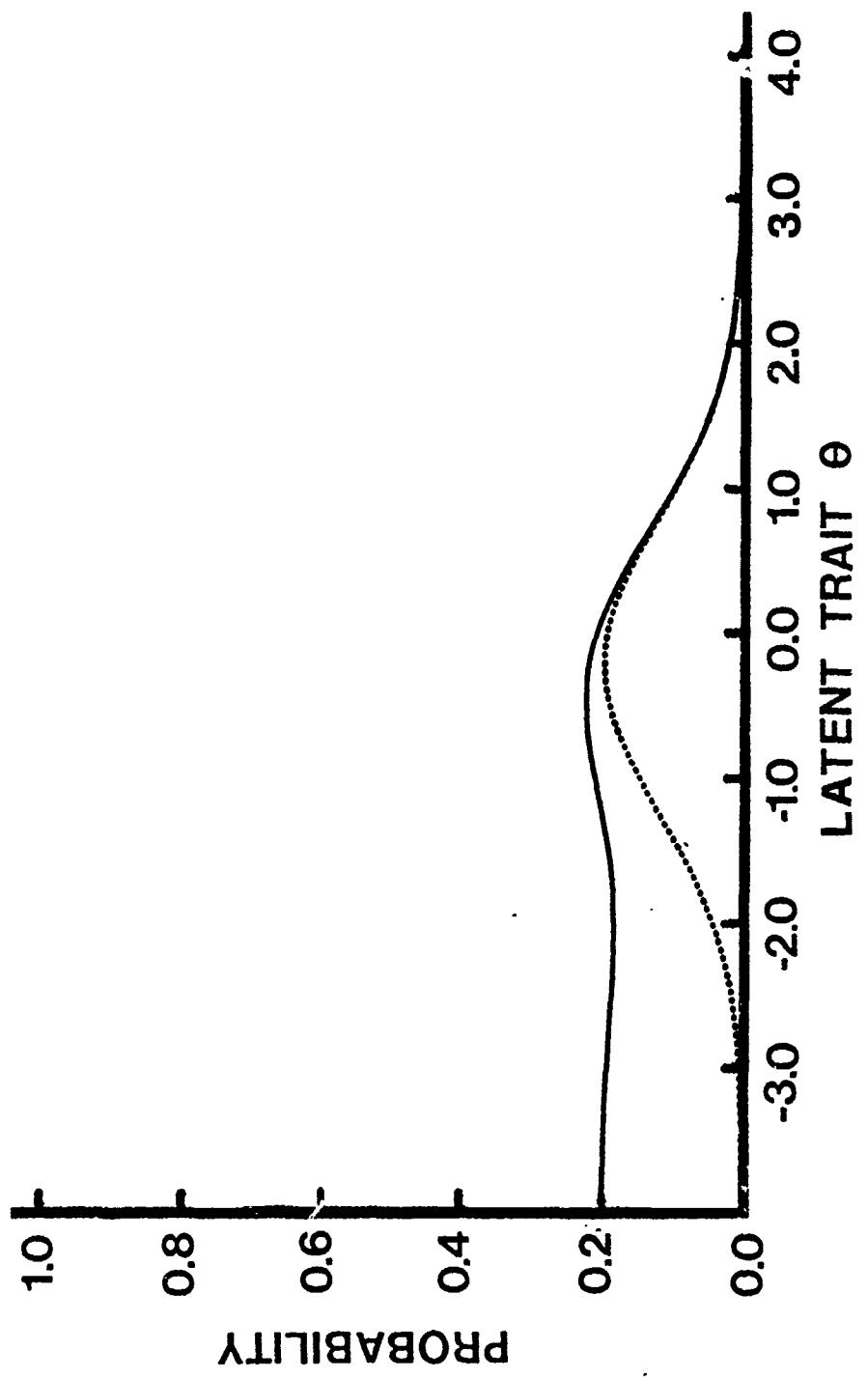


FIGURE 3-3 (Continued)  $h = 2$ .

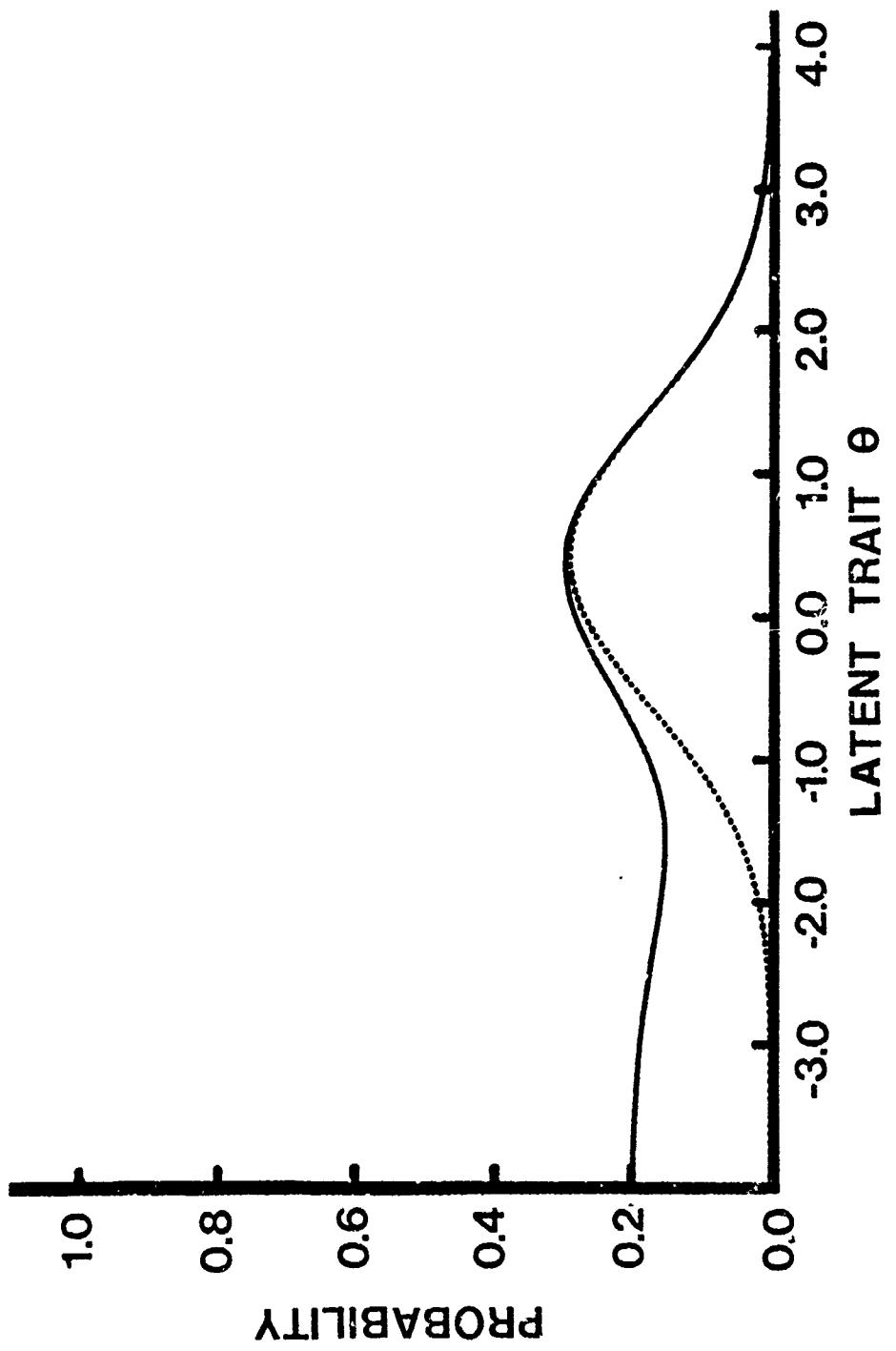


FIGURE 3-3 (Continued)  $h = 3$ .

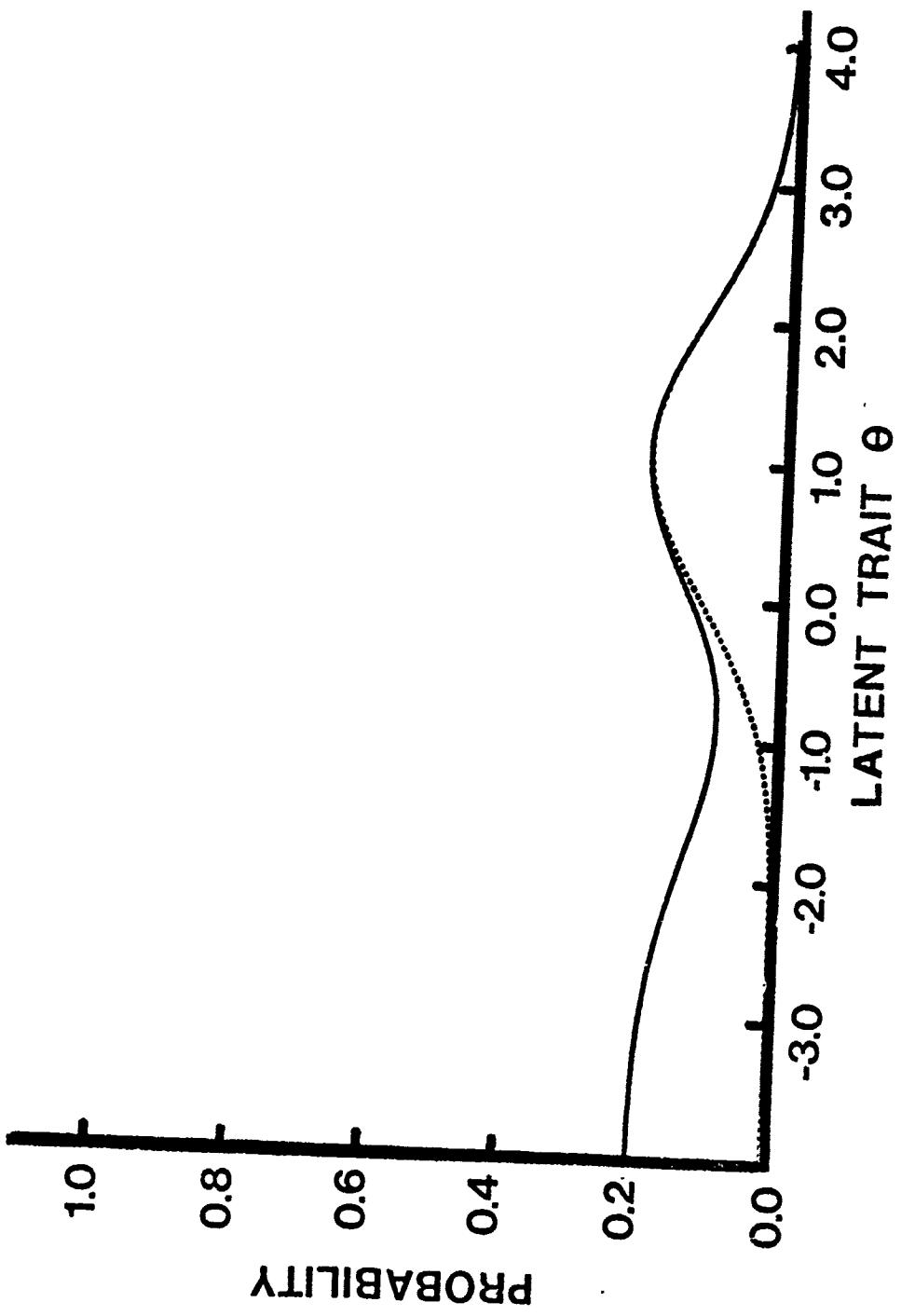


FIGURE 3-3 (Continued)  $h = 4$ .

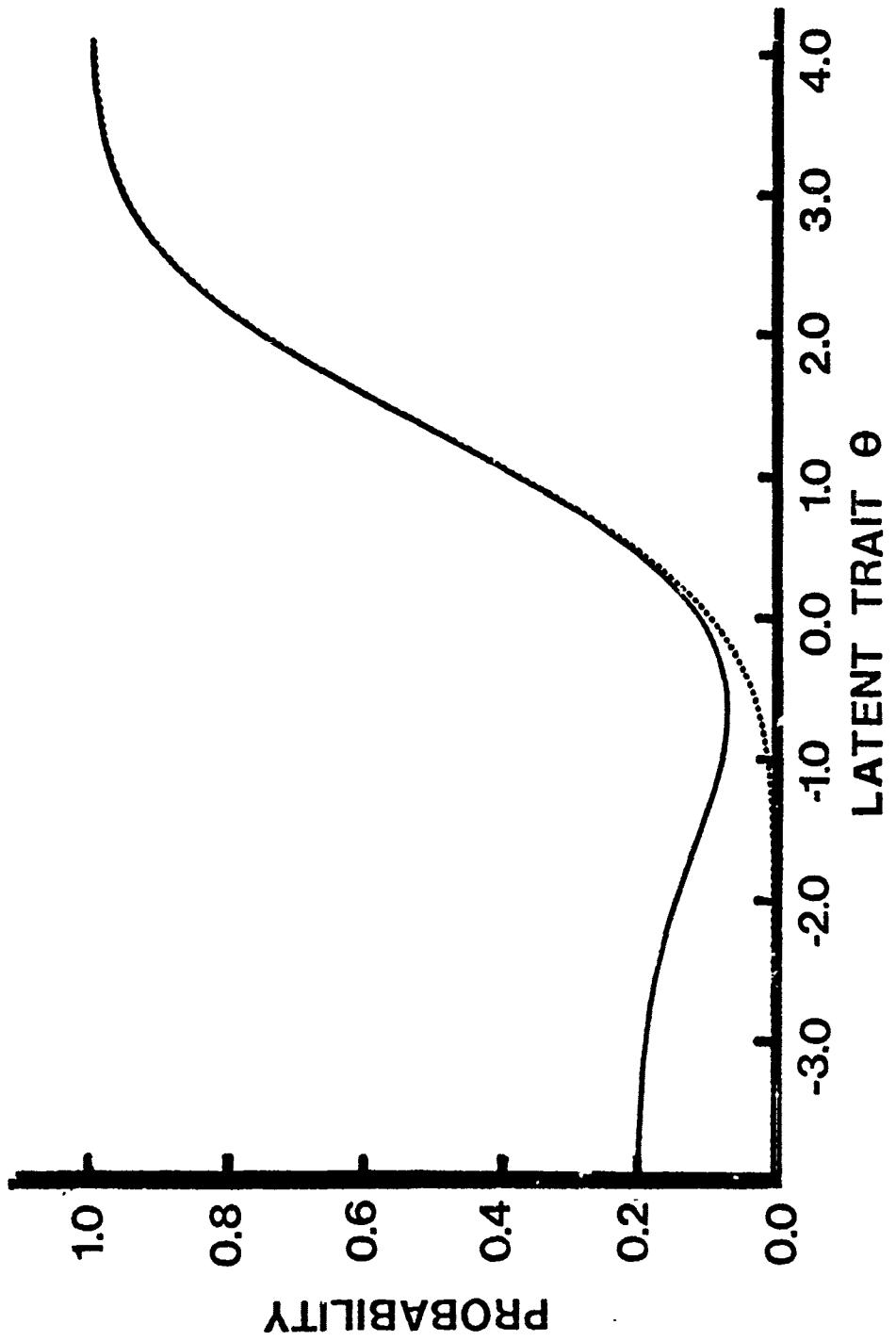


FIGURE 3-3 (Continued)  $h = 5$ .

in the case where  $b_h$ 's ( $h=1,2,\dots,m$ ) are close to one another the operating characteristic of the correct answer looks as if it followed the three-parameter normal ogive, or logistic, model with the lower asymptote less than  $1/m$ , which has been asserted often (e.g., Lord, 1968).

As another example, Figures 3-4 through 3-6 illustrate a similar set of graphs for an item whose parameters are:  $a_g = 2.00$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ . In contrast to the preceding example, except for  $h = 1$  and  $2$ , the operating characteristics given by Model A of Type I for the multiple-choice item are not so different from those of the normal ogive model on the graded response level, except for the "tails" which lie on the negative side of the ability dimension. This is mostly due to the higher value of the discrimination parameter  $a_g$  for this item. Two more sets of graphs of a similar nature are given in Appendix I, i.e., Figures A-1-1 through A-1-3 for an item with the parameters,  $a_g = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -1.00$ ,  $b_3 = -0.50$ ,  $b_4 = 0.00$  and  $b_5 = 0.50$ , and Figures A-1-4 through A-1-6 for an item with  $a_g = 2.50$ ,  $b_1 = -1.75$ ,  $b_2 = -0.75$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 1.75$ , both of which follow the normal ogive model on the graded response level and Model A of Type I for the multiple-choice item.

As an empirical example, which a model of Type I may fit, Table 3-1 presents a contingency table between the four alternatives of a multiple-choice test item, A, B, C and D, and the five ability levels of examinees, which was selected from a preliminary study

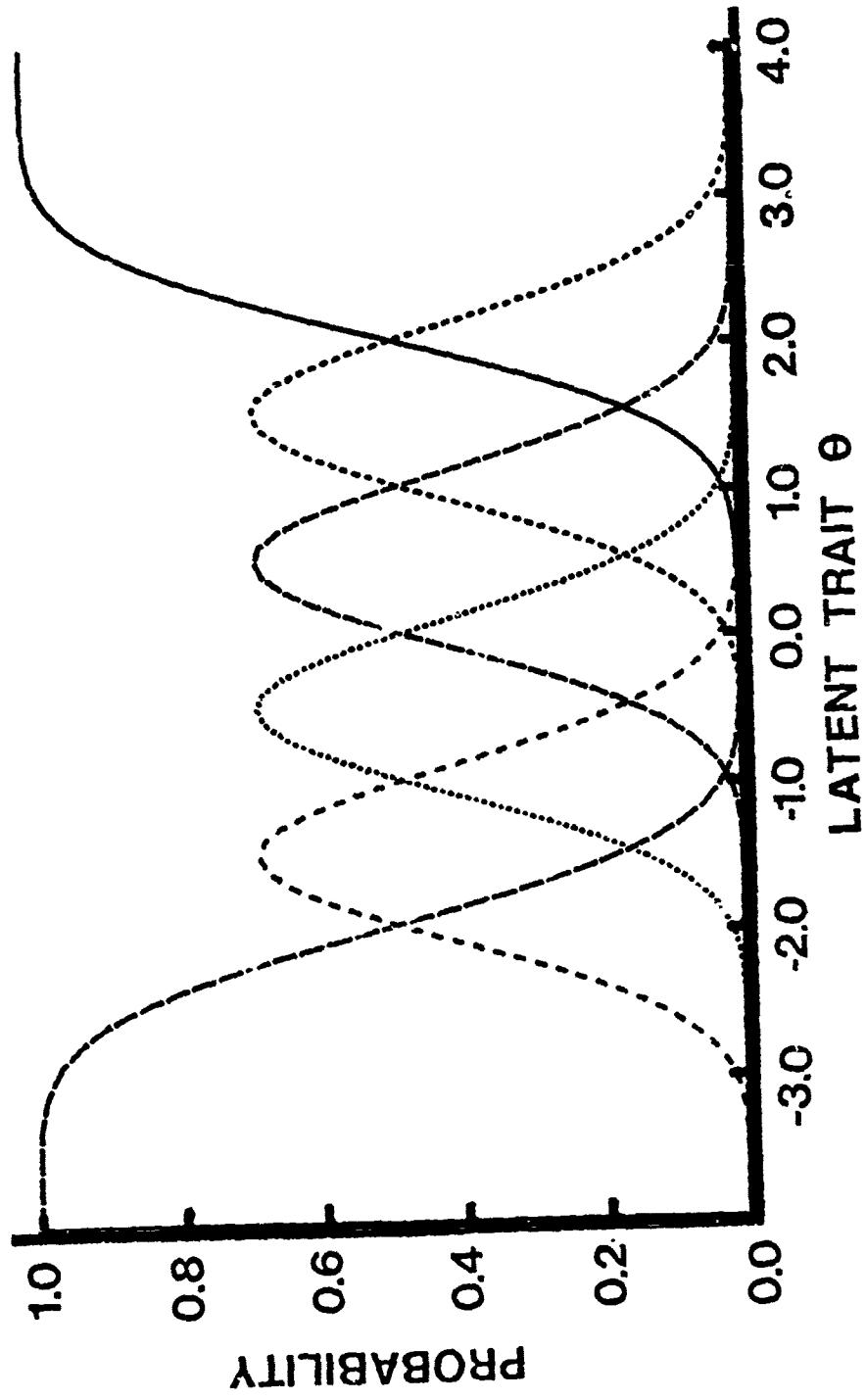


FIGURE 3-4  
Operating Characteristics of Six Item Response Categories Following the Normal Ogive Model, with  $a_g = 2.00$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .

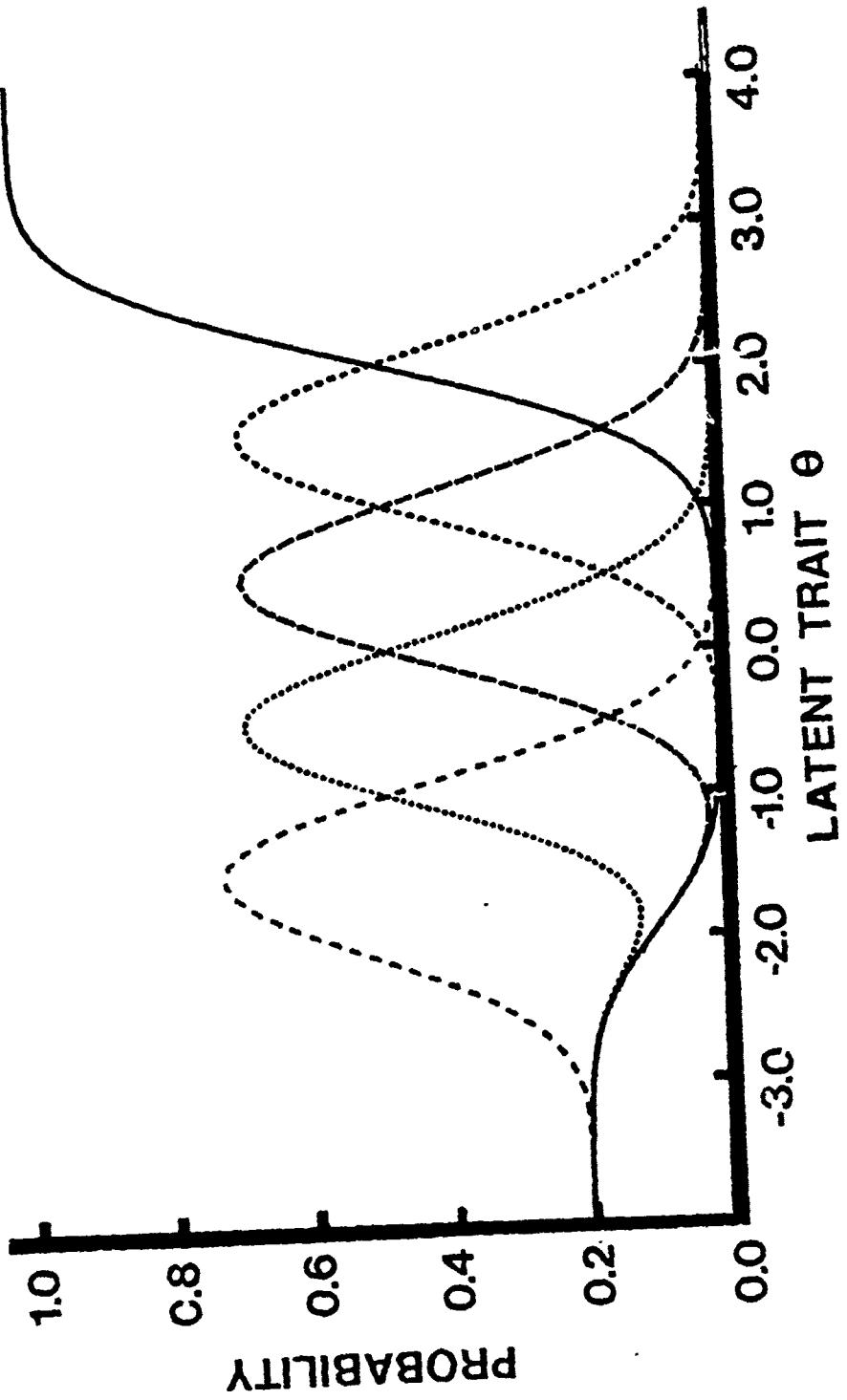


FIGURE 3-5

Operating Characteristics of Five Alternatives Following the Model A of Type I for the Multiple-Choice Item. The Parameters Are:  $a_g = 2.00$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .

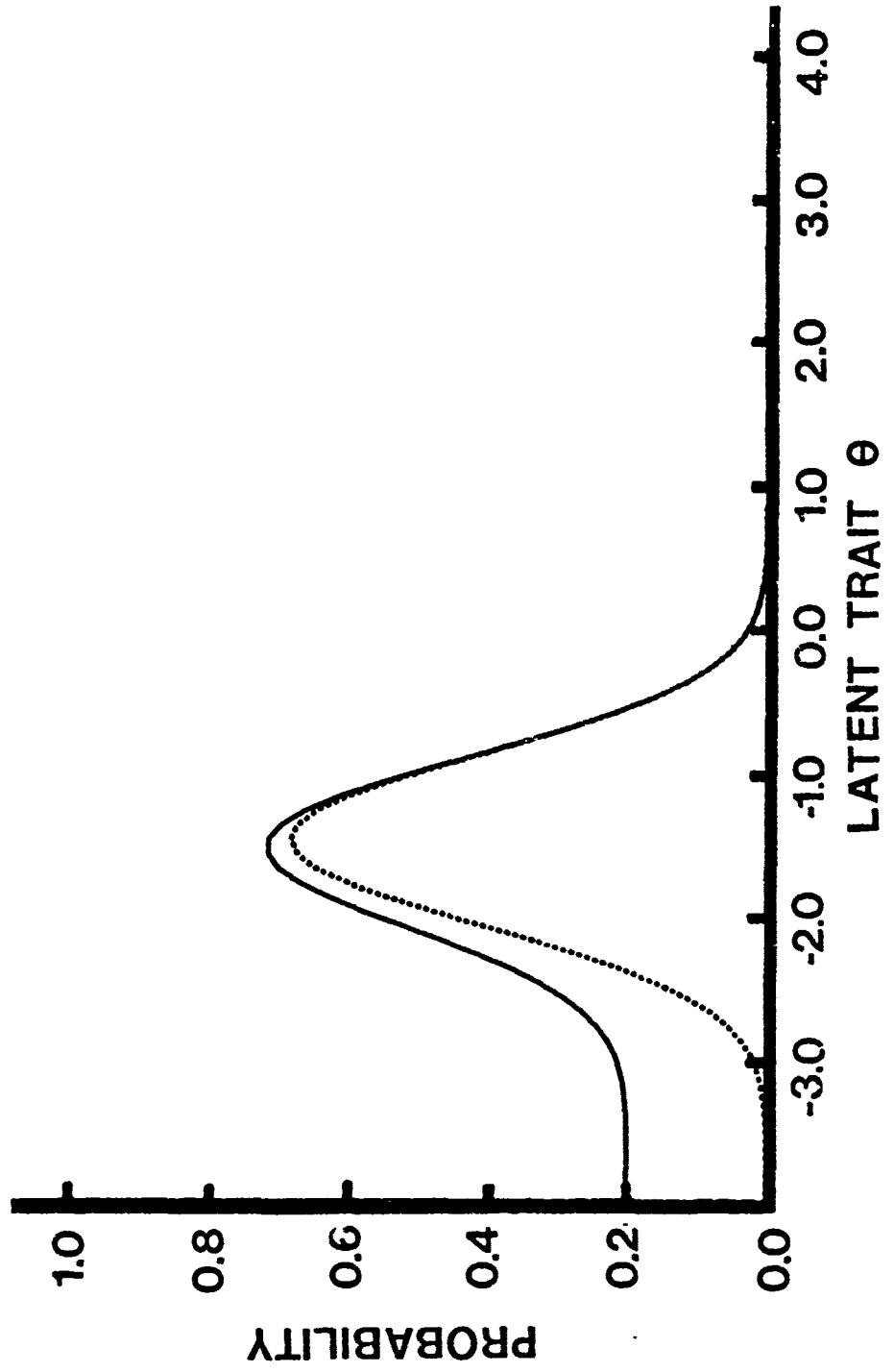


FIGURE 3-6

Comparison of the Two Operating Characteristics in the Normal Ogive Model  
(Dotted Curve) and in Model A of Type I (Solid Curve).  $h = 1$ .

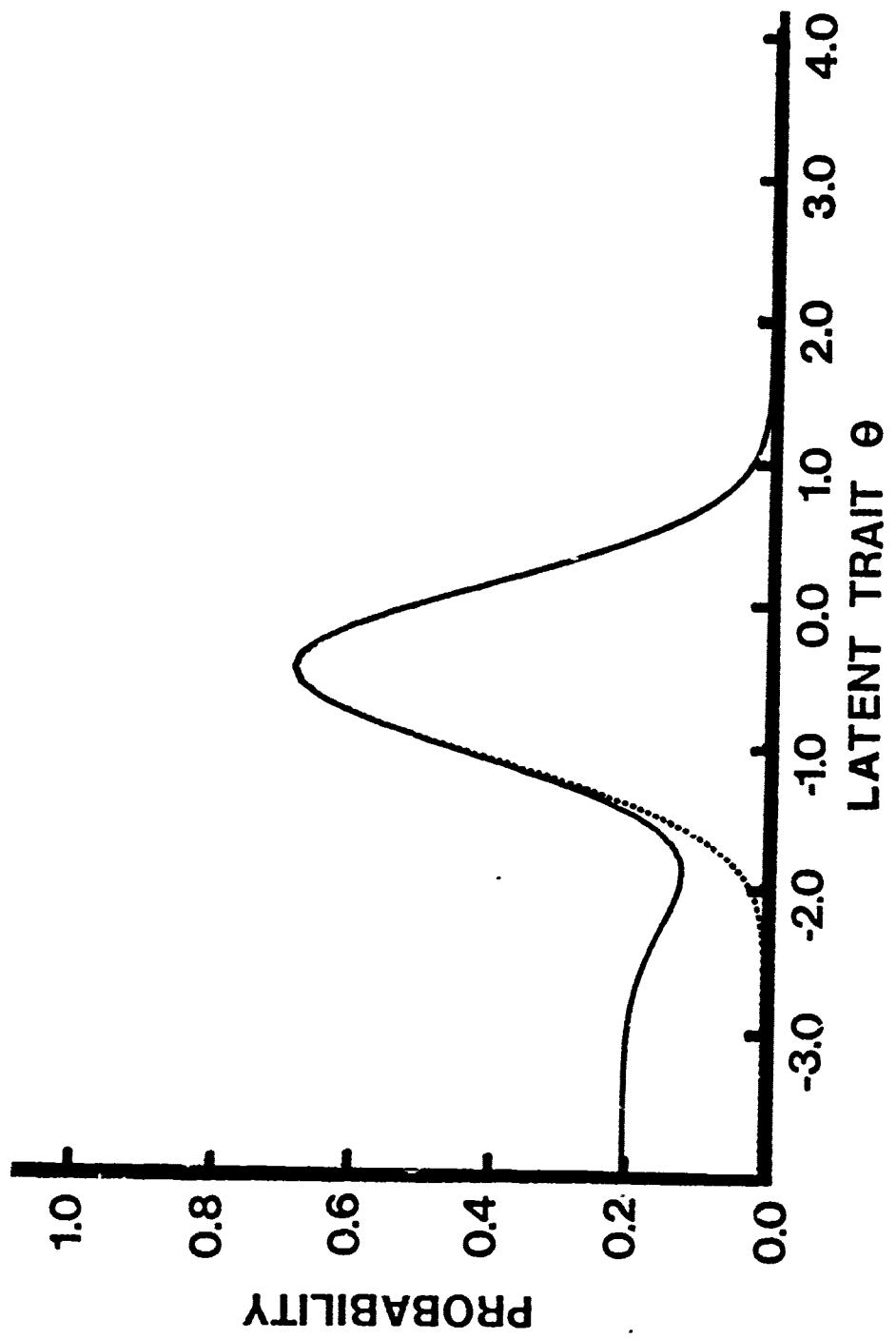


FIGURE 3-6 (Continued)  $h = 2$ .

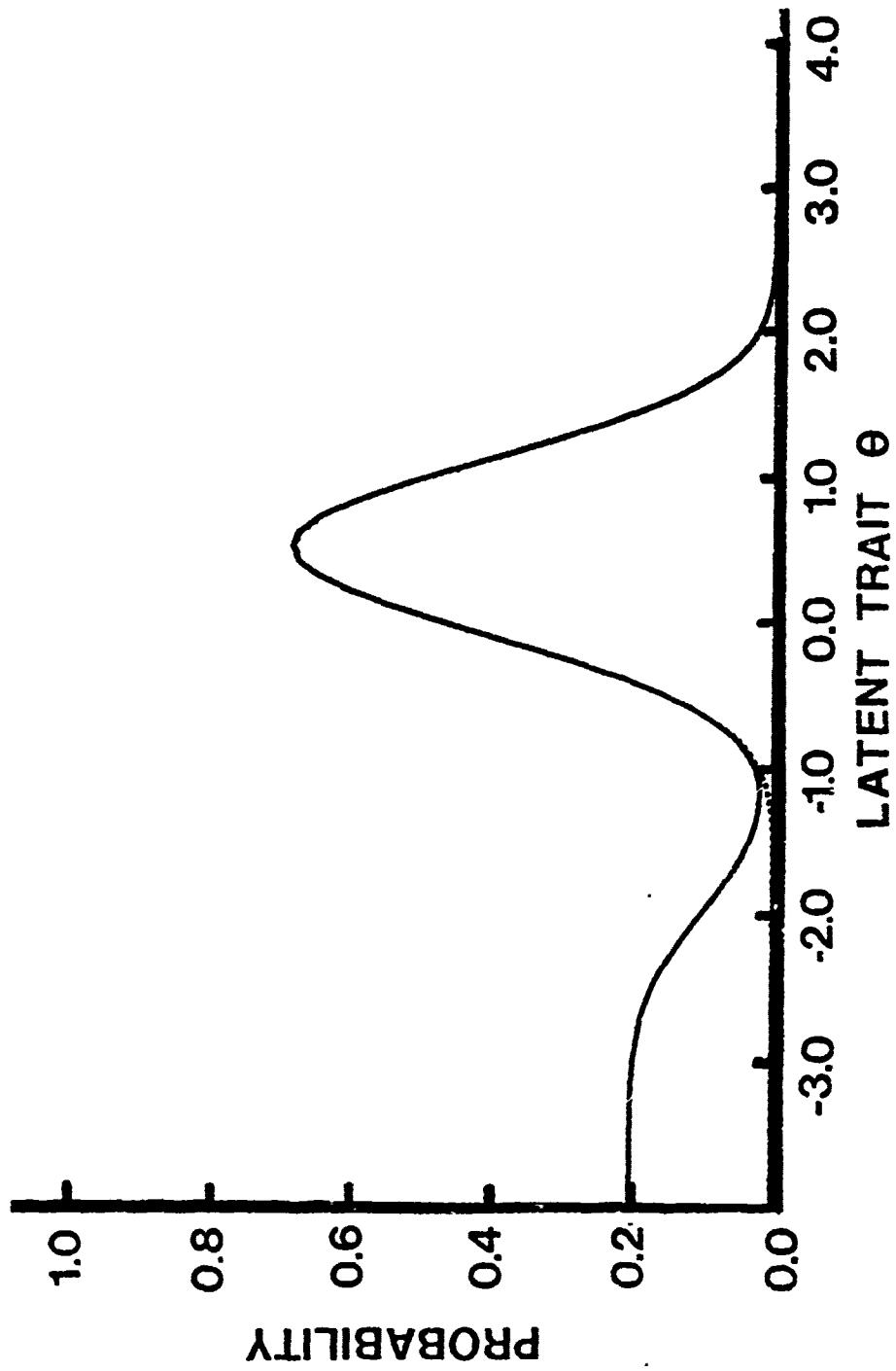


FIGURE 3-6 (Continued)  $h = 3$ .

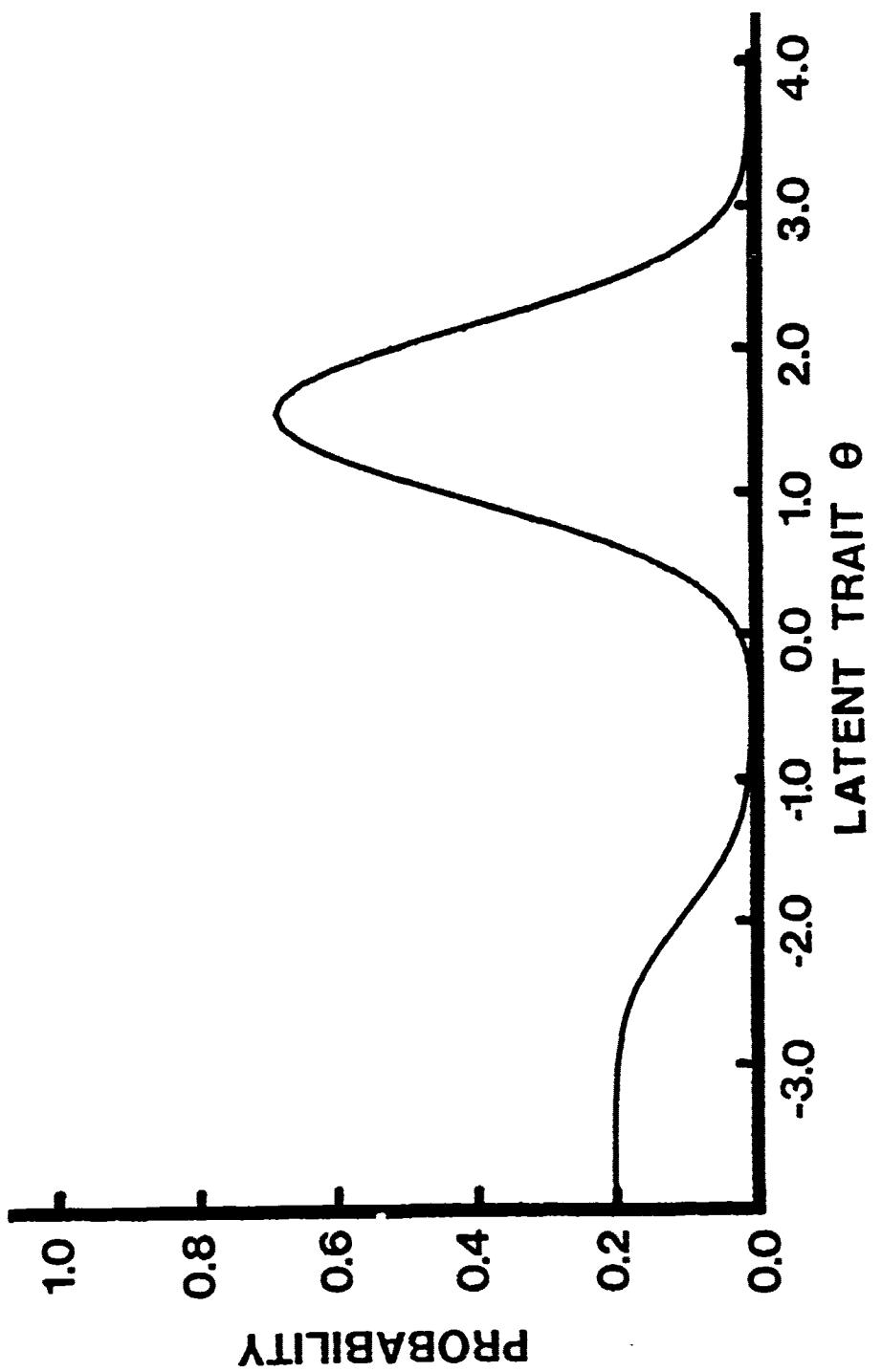


FIGURE 3-6 (Continued)  $h = 4$ .

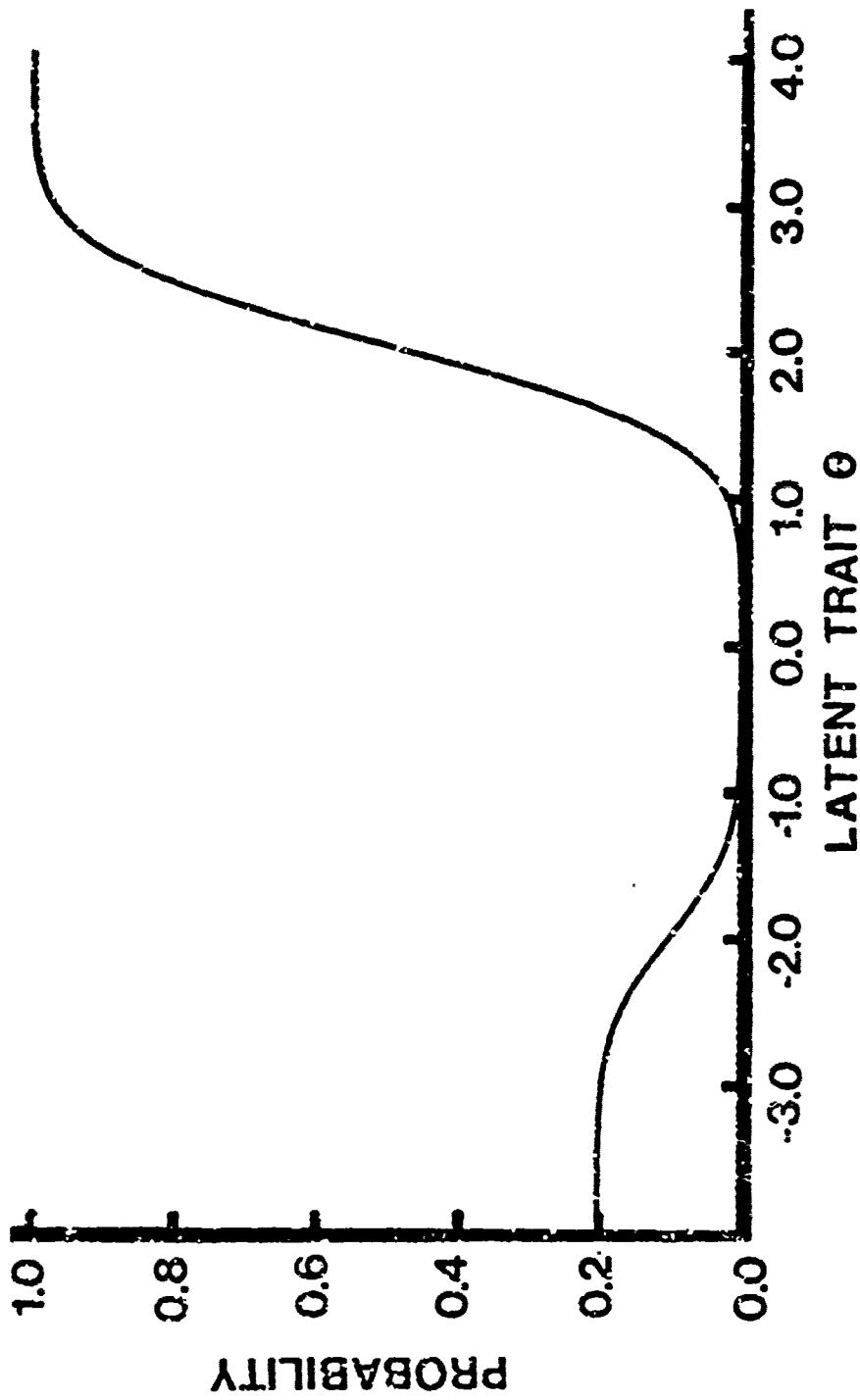


FIGURE 3-6 (Continued)  $h = 5$ .

TABLE 3-1  
Contingency Table Between the Four Alternatives and the Five  
Ability Groups for Item 43.

Alternative	Very Low	Low	Middle	High	Very High	Total
A	55	40	26	26	12	159
B	64	58	68	80	52	322
C	34	64	70	74	130	372
D	34	36	29	19	6	124
No Answer	1	0	0	0	0	1
Total	188	198	193	199	200	978

in test development.\* We notice that, for this multiple-choice test item, the mode of the frequency for the alternative A is the lowest ability group, that for the alternative B is the second highest ability group, that for C is the highest ability group, which is natural since this is the correct answer, and that for D is the second lowest ability group. Thus these categories may be ordered as A, D, B and C in the ascending order. Another similar example is shown in Table 3-2. For this item, the correct answer is the alternative C. The modes of the frequency for the alternatives, A, B, C and D, are the second lowest ability group, the highest ability group, the lowest ability group, and the lowest ability group, respectively. In this example, the order is not so clear for the alternatives, C and D, since both have similar frequency distributions.

Figures 3-7 through 3-11 present the operating characteristics of the six response categories in the logistic model, and those of the five alternatives in Model B, which follow the mathematical form given as (3.4), for five hypothetical multiple-choice items. In each figure, the discrimination parameter,  $a_g$ , and the difficulty parameters,  $b_h$  ( $h=1,2,3,4,5$ ), are specified. It can be seen that the set of parameters for item 2 is the same as the one used in Figures 3-1 through 3-3 for the normal ogive model and Model A, and the set of parameters for item 4 is identical with the one used in Figures 3-4 through 3-6. As is expected, the resulting curves are

---

\*The author is obliged to Mr. Lon Raske and Educational Testing Service for allowing her to use their data, which include this and the following examples.

TABLE 3-2  
Contingency Table Between the Four Alternatives and the Five  
Ability Groups for Item 46.

Alternative	Very Low	Low	Middle	High	Very High	Total
A	70	104	89	69	67	399
B	31	28	54	99	121	333
C	42	31	28	15	7	123
D	43	32	20	11	4	110
No Answer	0	2	0	0	0	2
Total	186	197	191	194	199	967



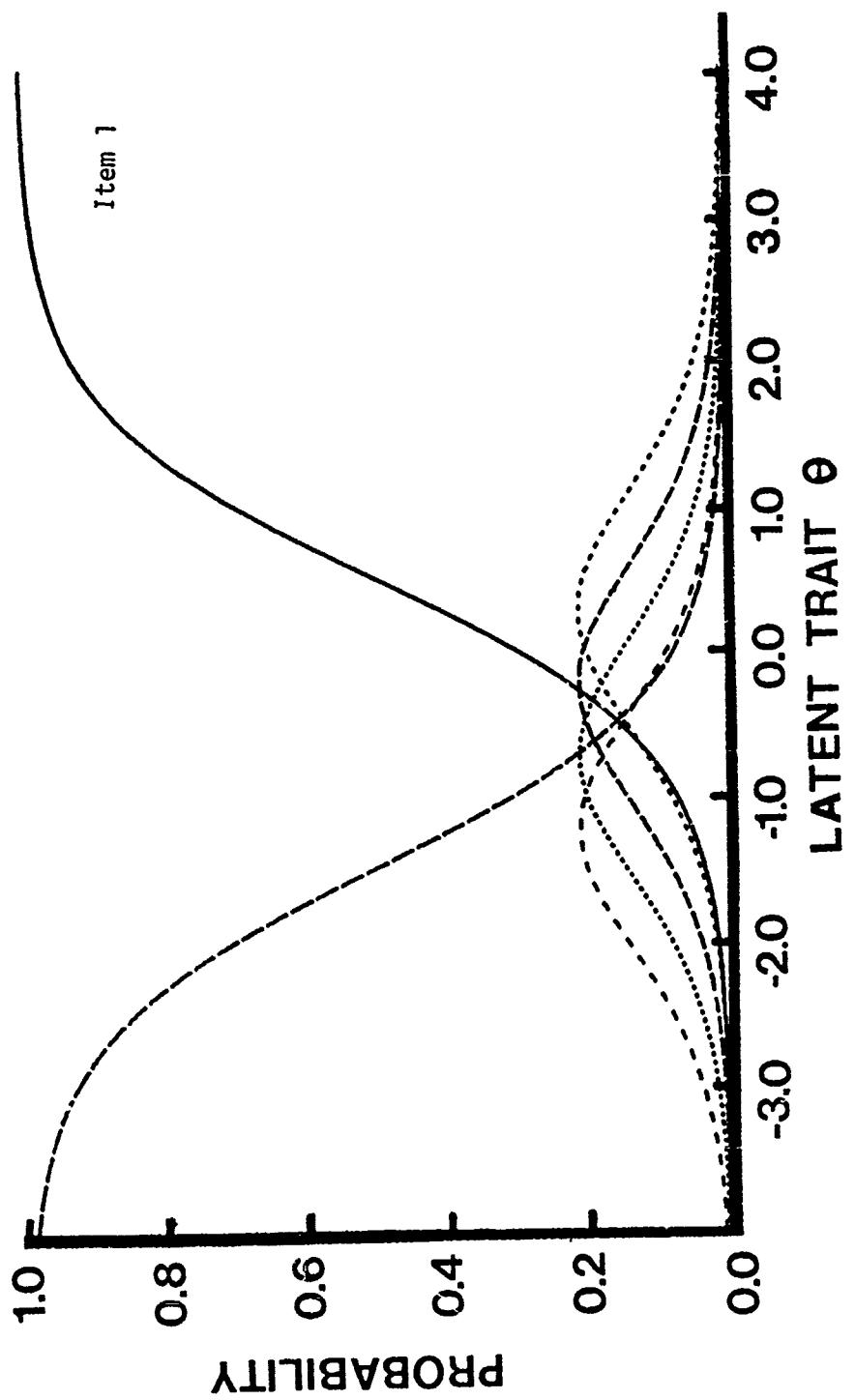


FIGURE 3-7

Operating Characteristics of Six Item Response Categories Following the Logistic Model, with  $a_g = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -1.00$ ,  $b_3 = -0.50$ ,  $b_4 = 0.00$  and  $b_5 = 0.50$ .

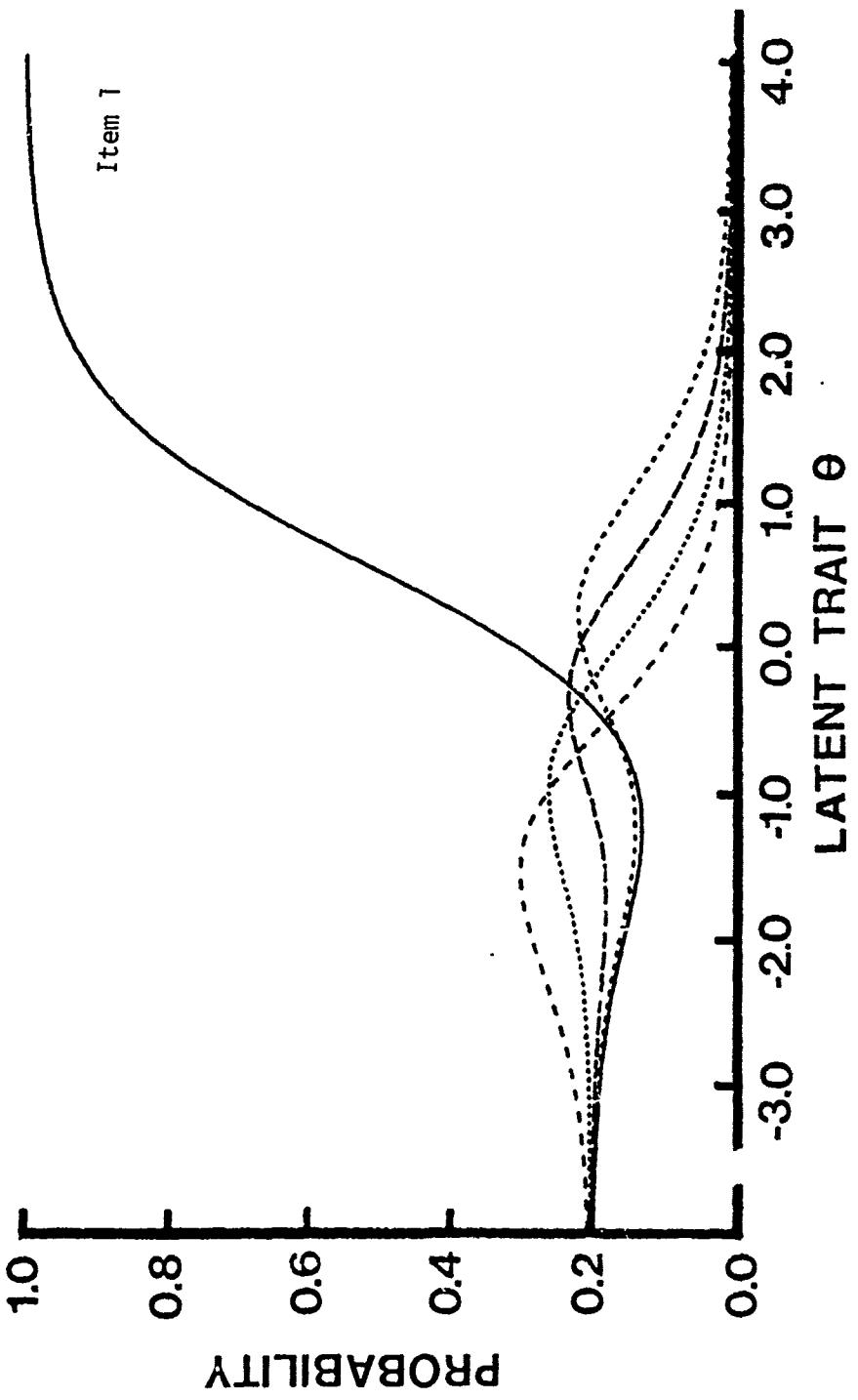


FIGURE 3-7 (Continued)

Operating Characteristics of Five Alternatives Following Model B of Type I for the Multiple-Choice Item, with the Same Parameters.

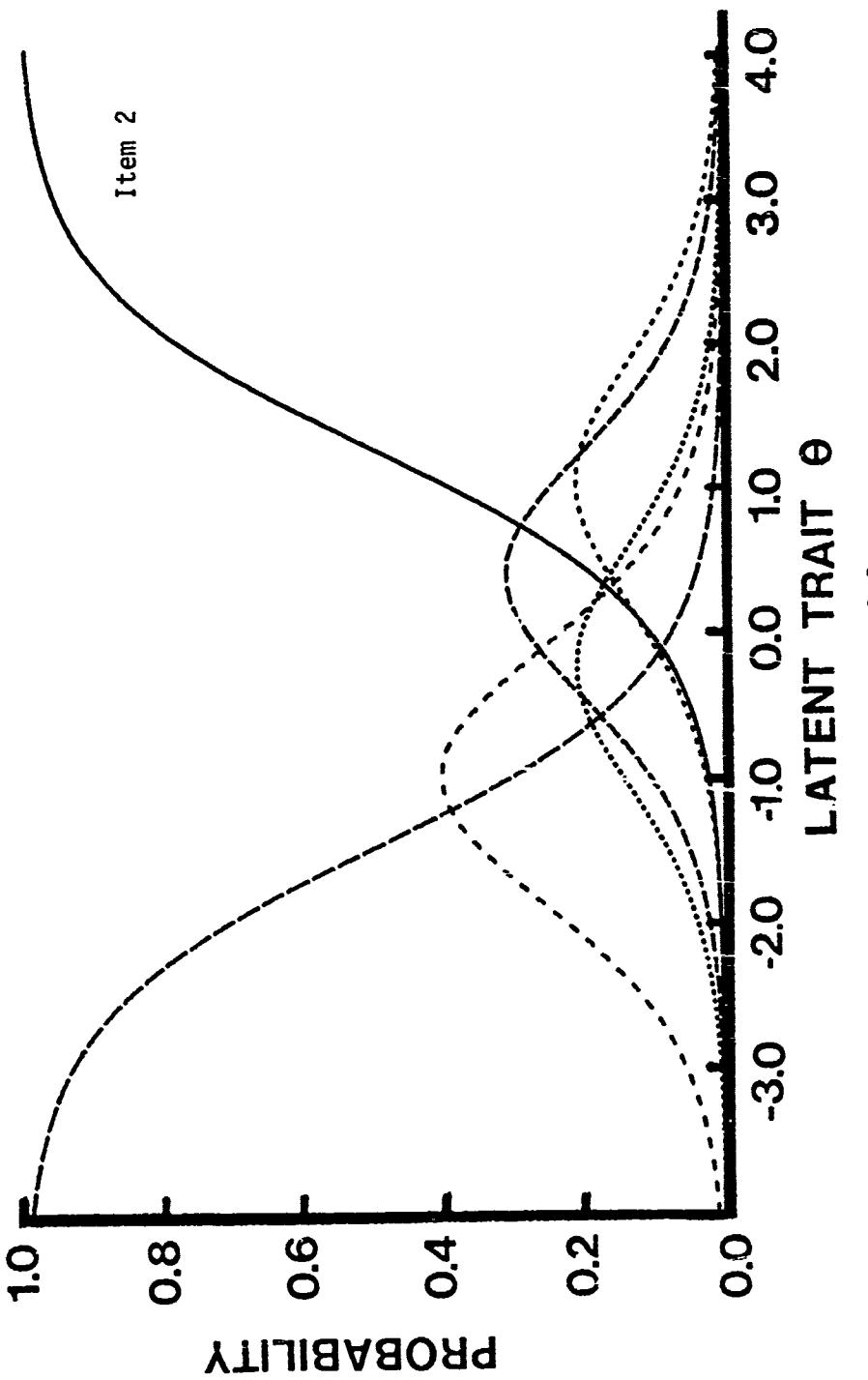


FIGURE 3-8

Operating Characteristics of Six Item Response Categories Following the Logistic Model, with  $a_g = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -0.50$ ,  $b_3 = 0.00$ ,  $b_4 = 0.75$  and  $b_5 = 1.25$ .

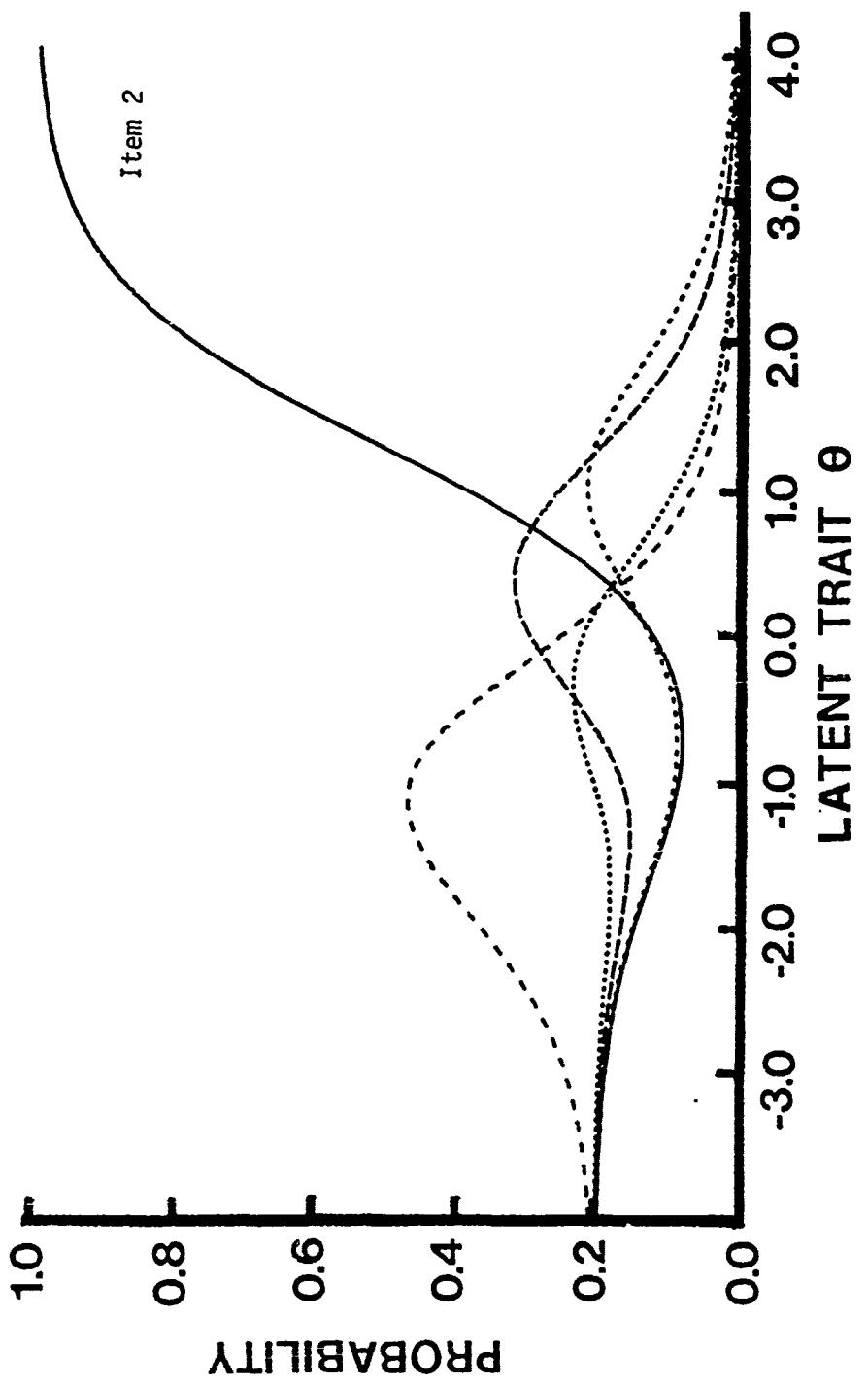


FIGURE 3-8 (Continued)

Operating Characteristics of Five Alternatives Following Model B of Type I for the Multiple-Choice Item, with the Same Parameters.

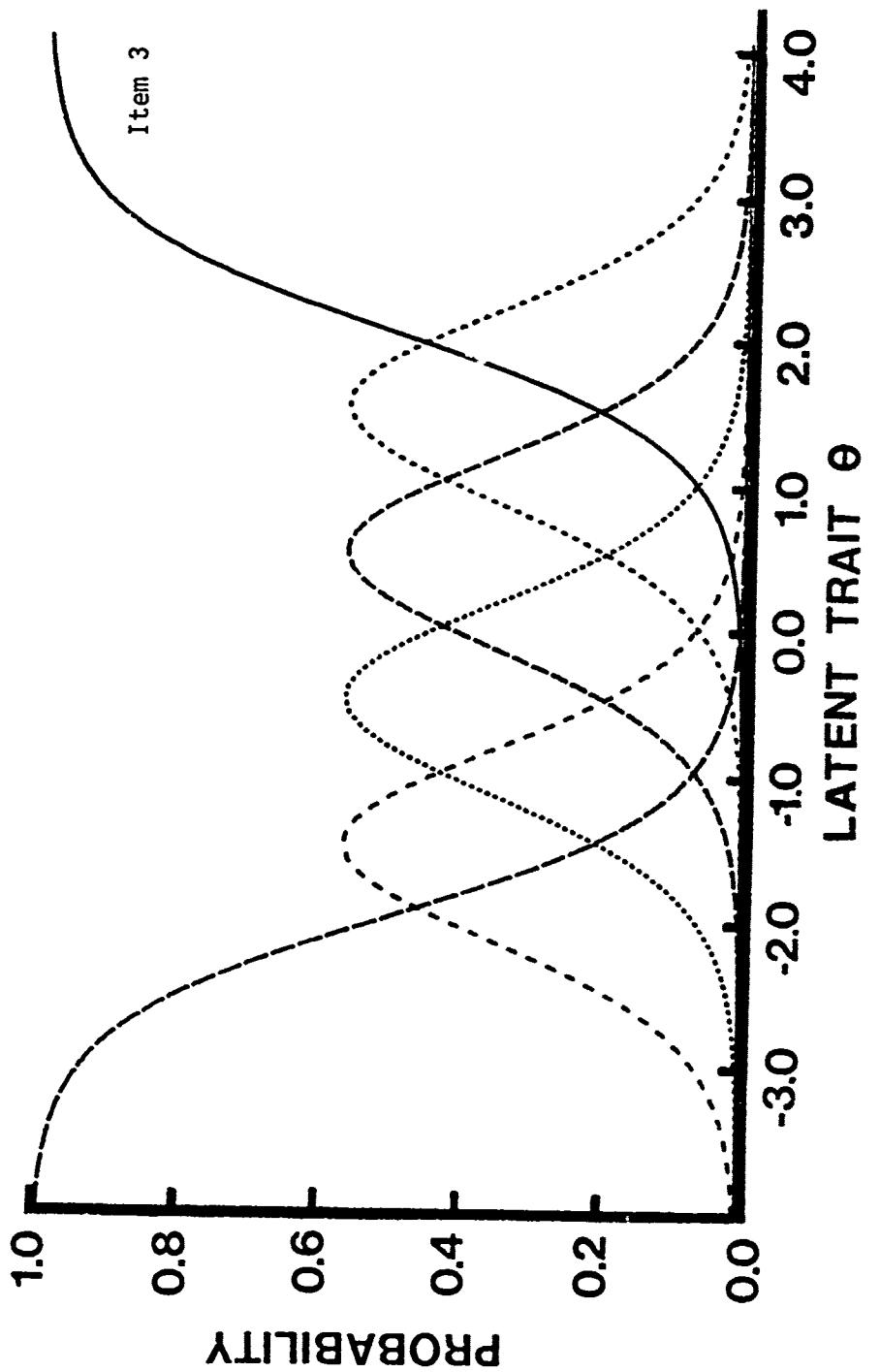


FIGURE 3-9

Operating Characteristics of Six Item Response Categories Following the Logistic Model, with  $a_8 = 1.50$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .

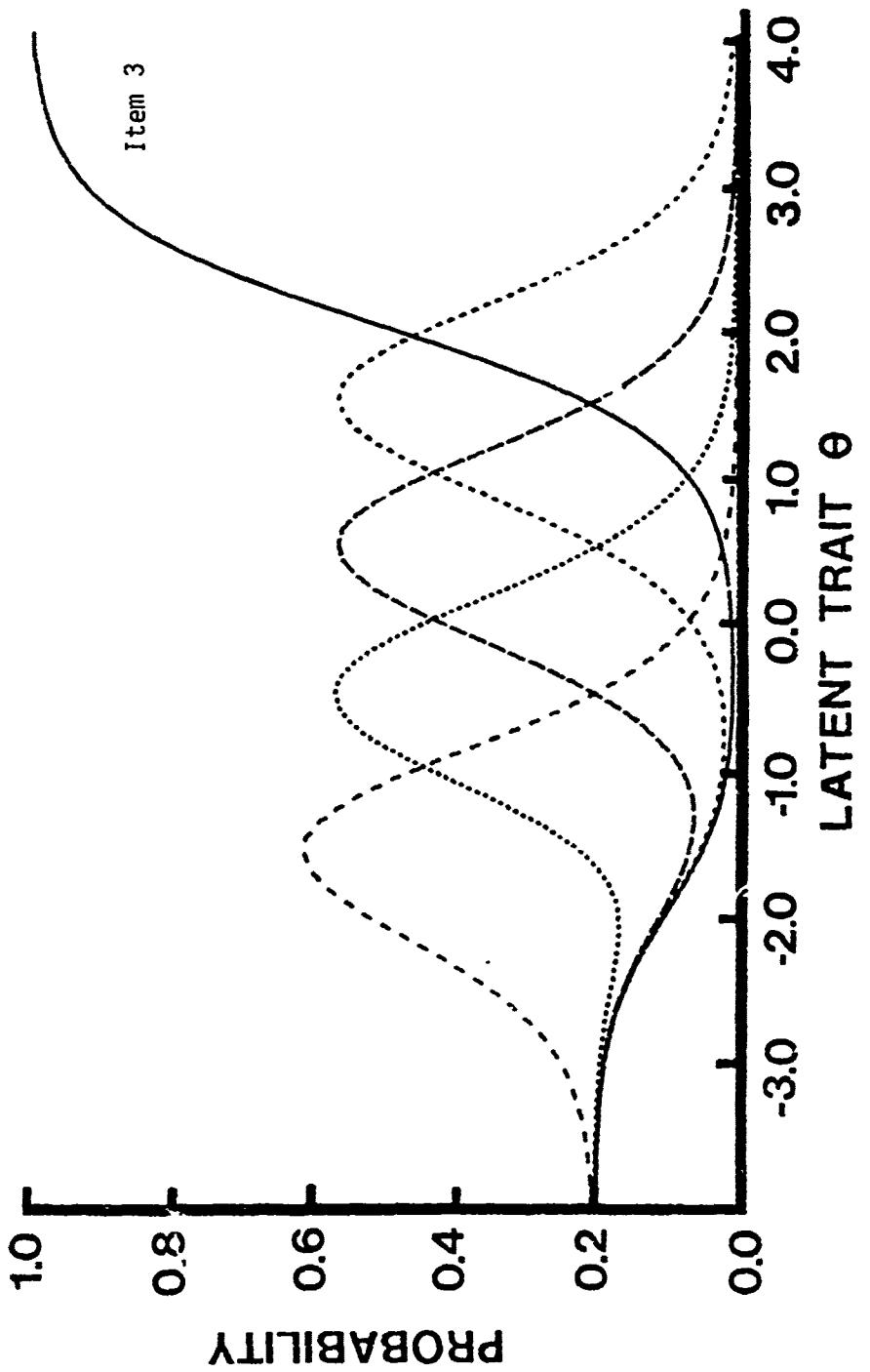


FIGURE 3-9 (Continued)

Operating Characteristics of Five Alternatives Following Model B of Type I for the Multiple-Choice Item, with the Same Parameters.

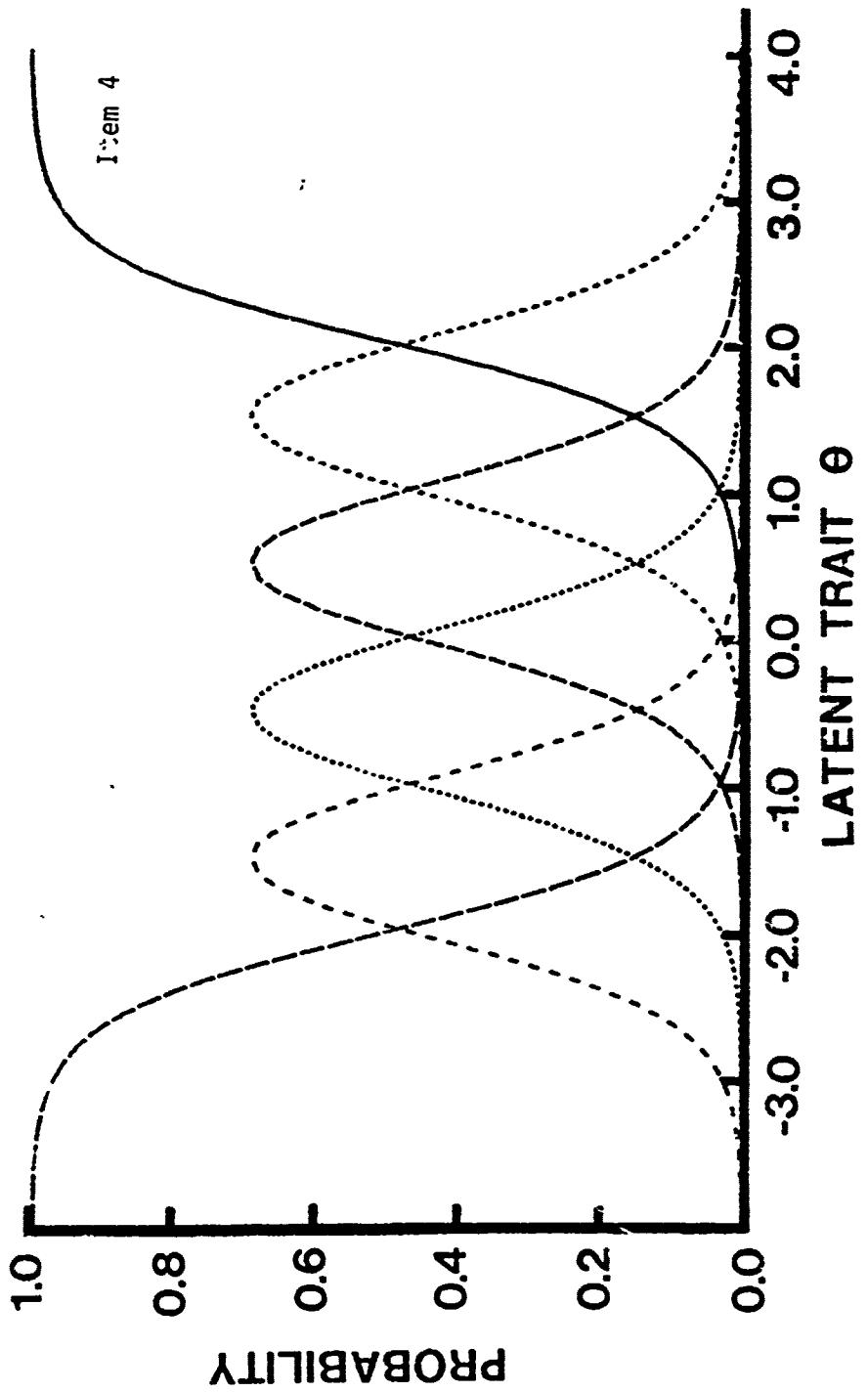


FIGURE 3-10

Operating Characteristics of Six Item Response Categories Following the Logistic Model, with  $a_g = 2.00$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .

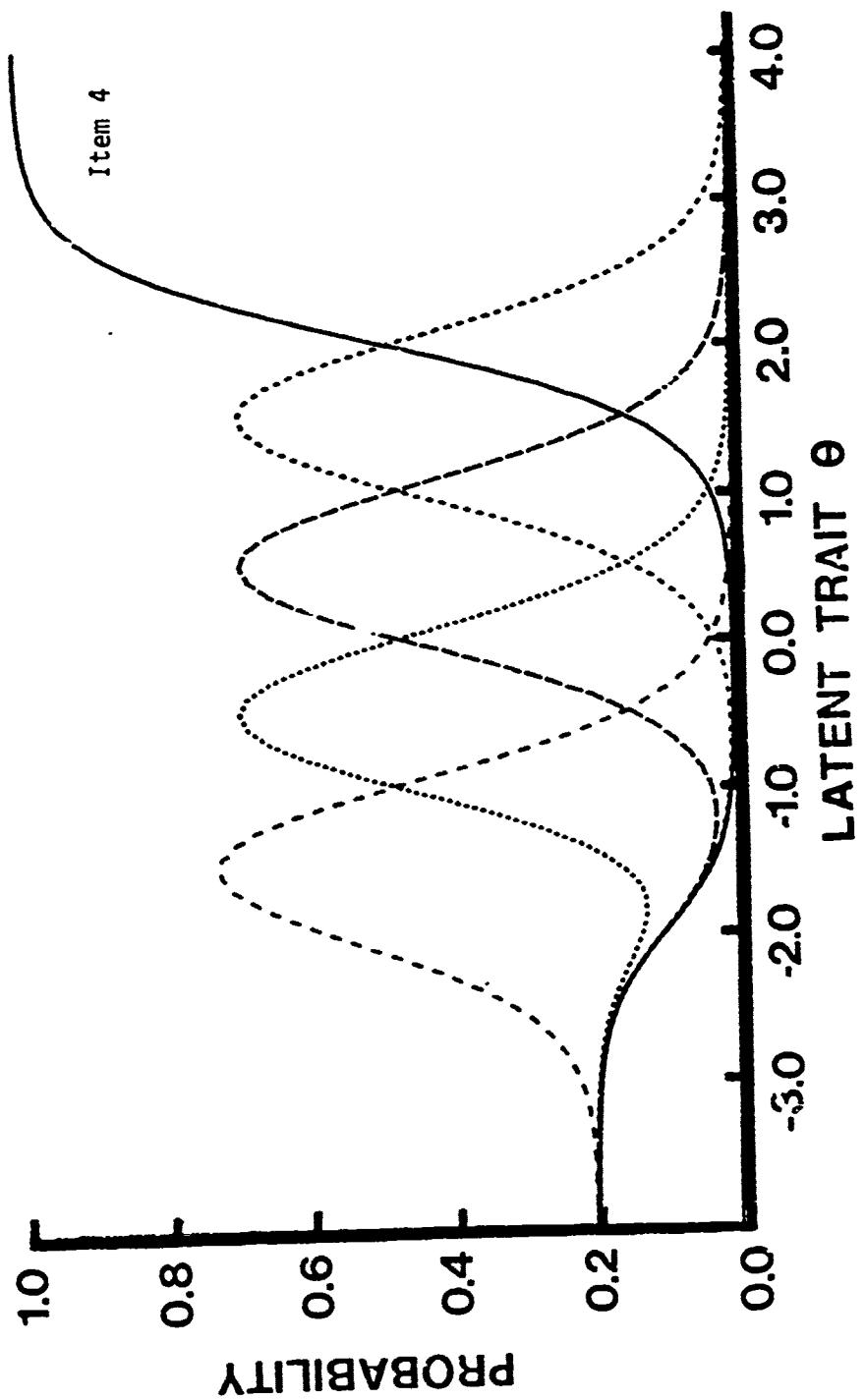


FIGURE 3-10 (Continued)

Operating Characteristics of Five Alternatives Following Model B of Type I for the Multiple-Choice Item, with the Same Parameters.

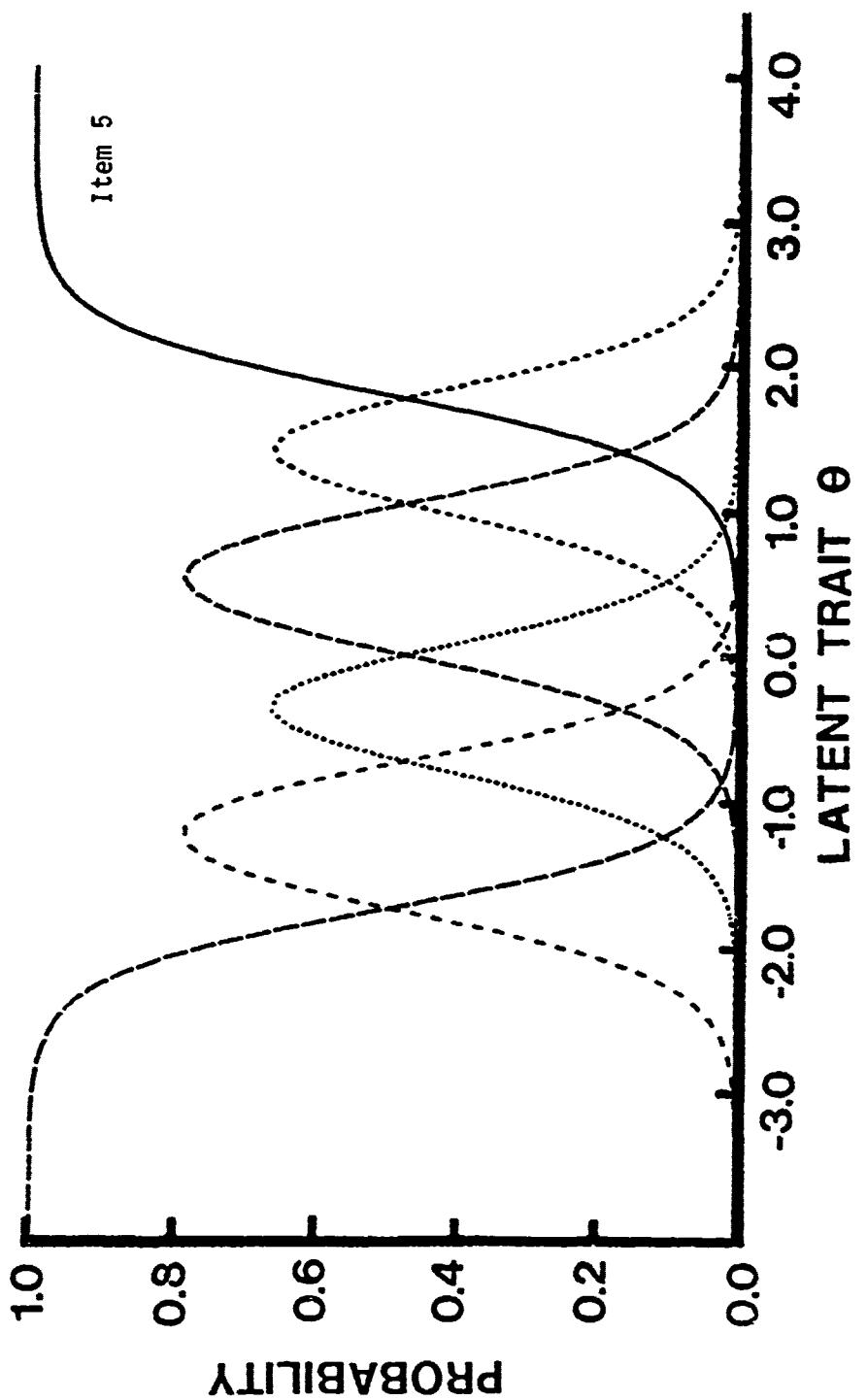


FIGURE 3-11

Operating Characteristics of Six Item Response Categories Following the Logistic Model, with  $a_8 = 2.50$ ,  $b_1 = -1.75$ ,  $b_2 = -0.75$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 1.75$ .

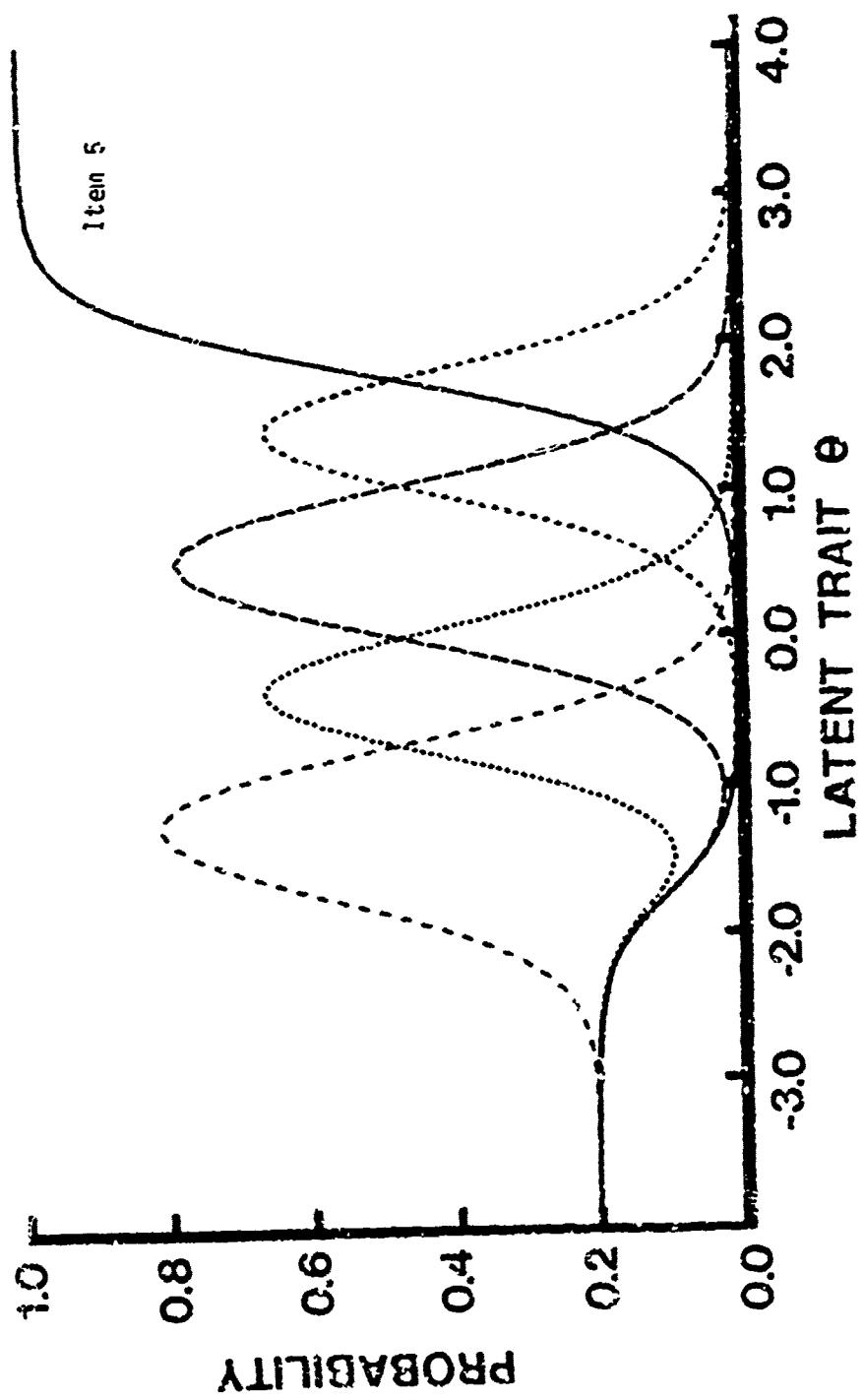


FIGURE 3-11 (Continued)

Operating Characteristics of Five Alternatives Following Model R of Type I for the Multiple-Choice Item, with the Same Parameters.

very similar between the normal ogive model and the logistic model, and between Models A and B. Identification of the curves with the categories can be made in the upper graph of each of Figures 3-7 through 3-11, since the modal points of these curves are in the ascending order of categories (cf. Samejima, 1972).

Figure 3-12 through 3-17 present the operating characteristics of the five response categories following the multinomial model, and those of the four alternatives in Model C, which are given by (3.5), for another set of six hypothetical multiple-choice test items. In each figure, the parameters  $a_h$  and  $c_h$  ( $h=1,2,3,4,5$ ) are specified. Identification of the curves with the categories can be done in the upper graph of each of Figures 3-12 through 3-17, since the modal points of these curves are in the ascending order of the values of  $a_h$  (Samejima, 1972). In the lower graphs of these figures, i.e., those for Model C, the asymptote at  $\theta \rightarrow -\infty$  is 0.25, instead of 0.20, since the number of alternatives in these examples is four, instead of five.



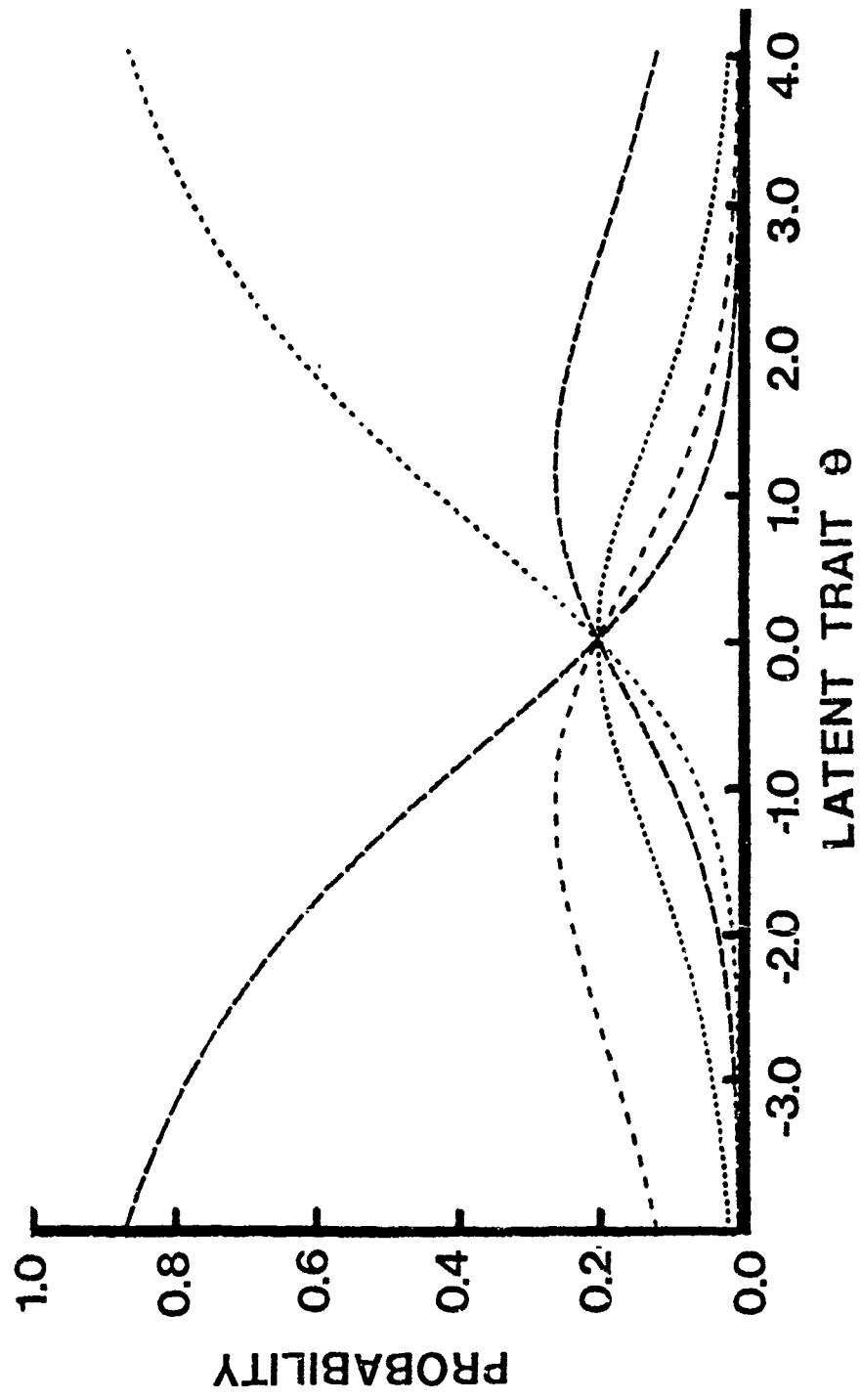


FIGURE 3-7.2

Operating Characteristics of Five Item Response Categories Following the Multinomial Model. The Item Parameters Are:  $a_1 = -1.000$ ,  $a_2 = -0.500$ ,  $a_3 = 0.000$ ,  $a_4 = 0.500$ ,  $a_5 = 1.000$ ;  $c_1 = 0.000$ ,  $c_2 = 0.000$ ,  $c_3 = 0.000$ ,  $c_4 = 0.000$ ,  $c_5 = 0.000$ .

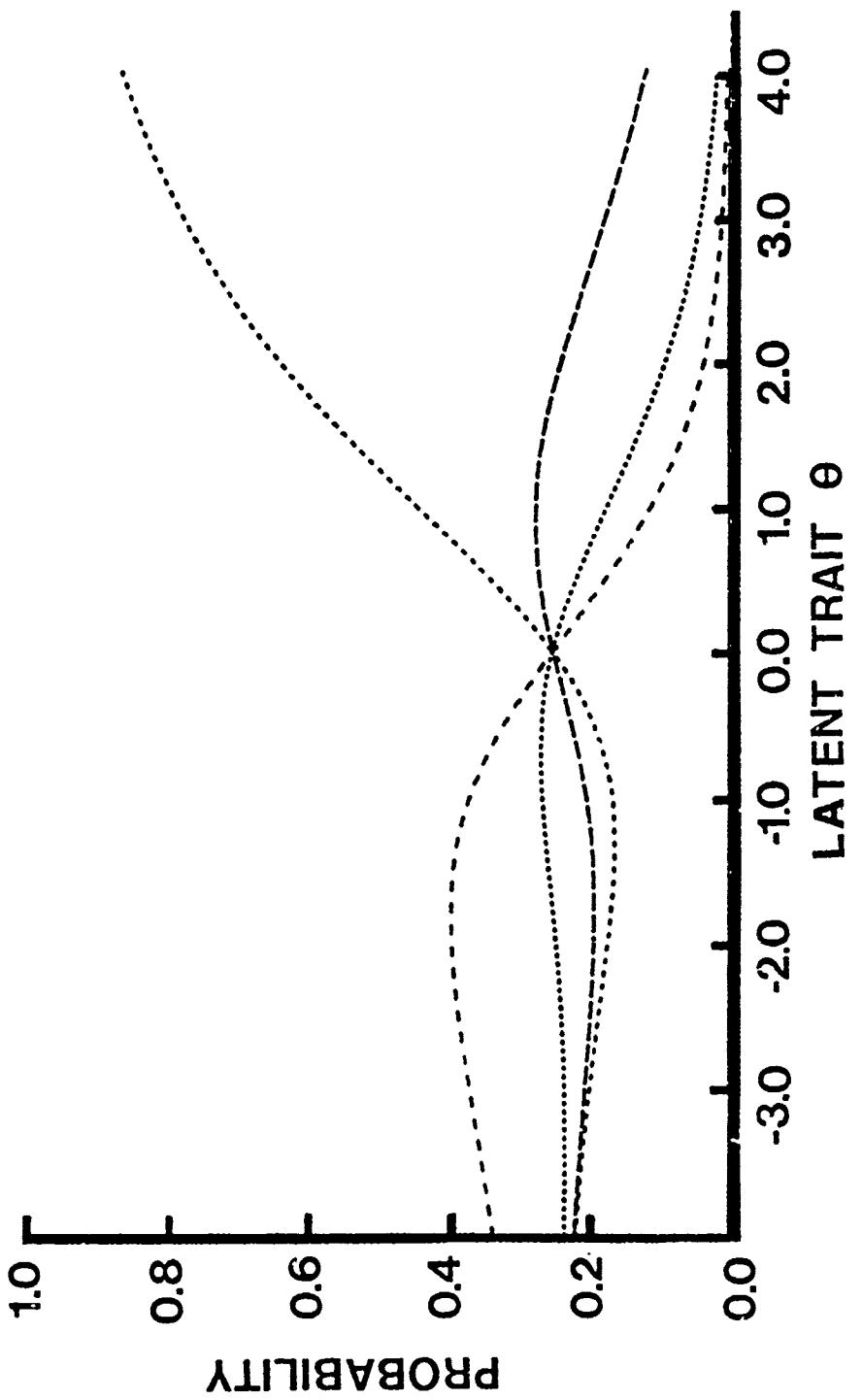


FIGURE 3-12 (Continued)

Operating Characteristics of Four Alternatives Following Model C of Type I  
for the Multiple-Choice Item, with the Same Parameters.

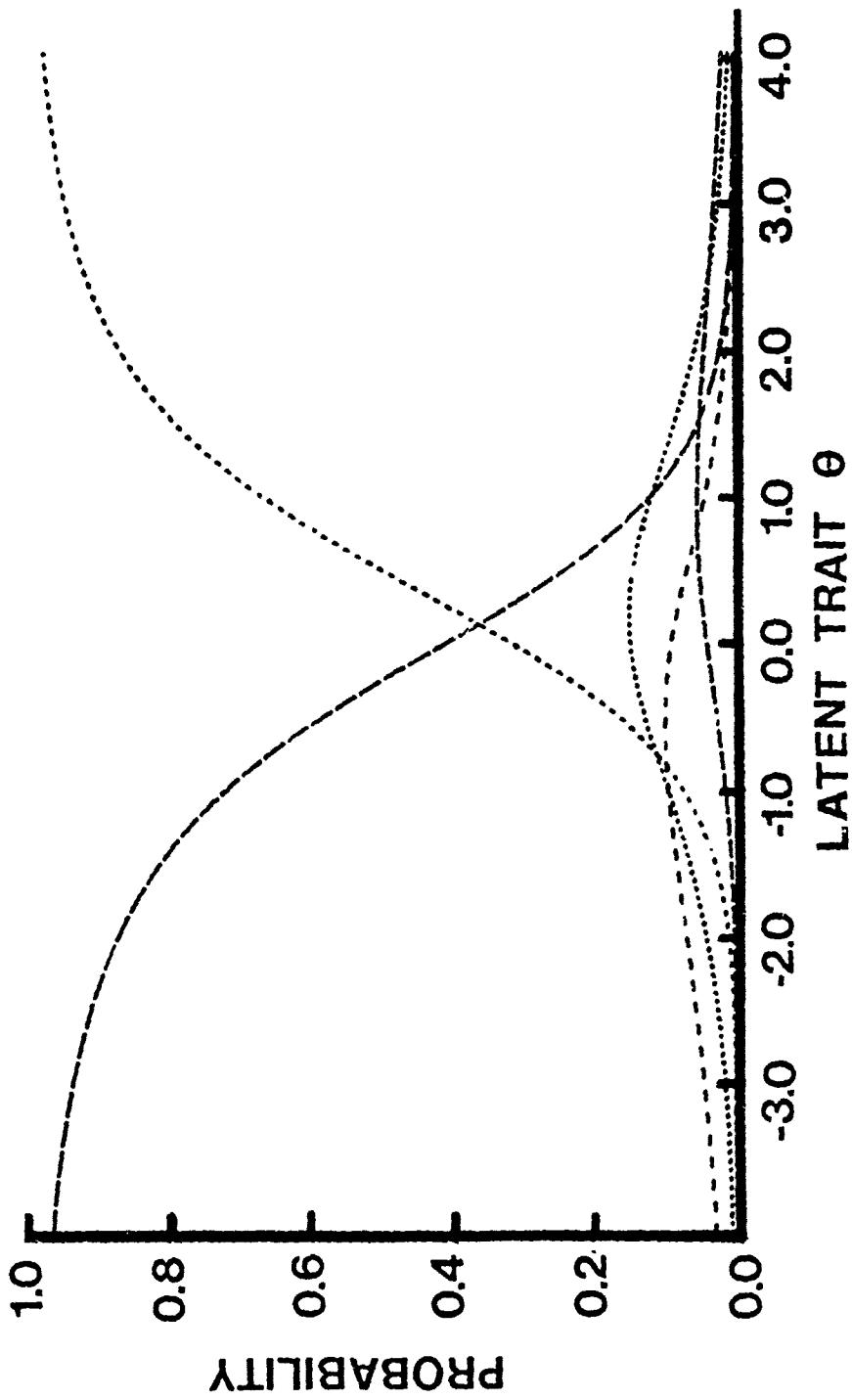


FIGURE 3-13

Operating Characteristics of Five Item Response Categories Following the Multinomial Model. The Item Parameters Are:  $a_1 = -1.000$ ,  $a_2 = -0.500$ ,  $a_3 = 0.000$ ,  $a_4 = 0.500$ ,  $a_5 = 1.000$ ;  $c_1 = 1.000$ ,  $c_2 = -0.500$ ,  $c_3 = 0.000$ ,  $c_4 = -1.250$ ,  $c_5 = 0.750$ .

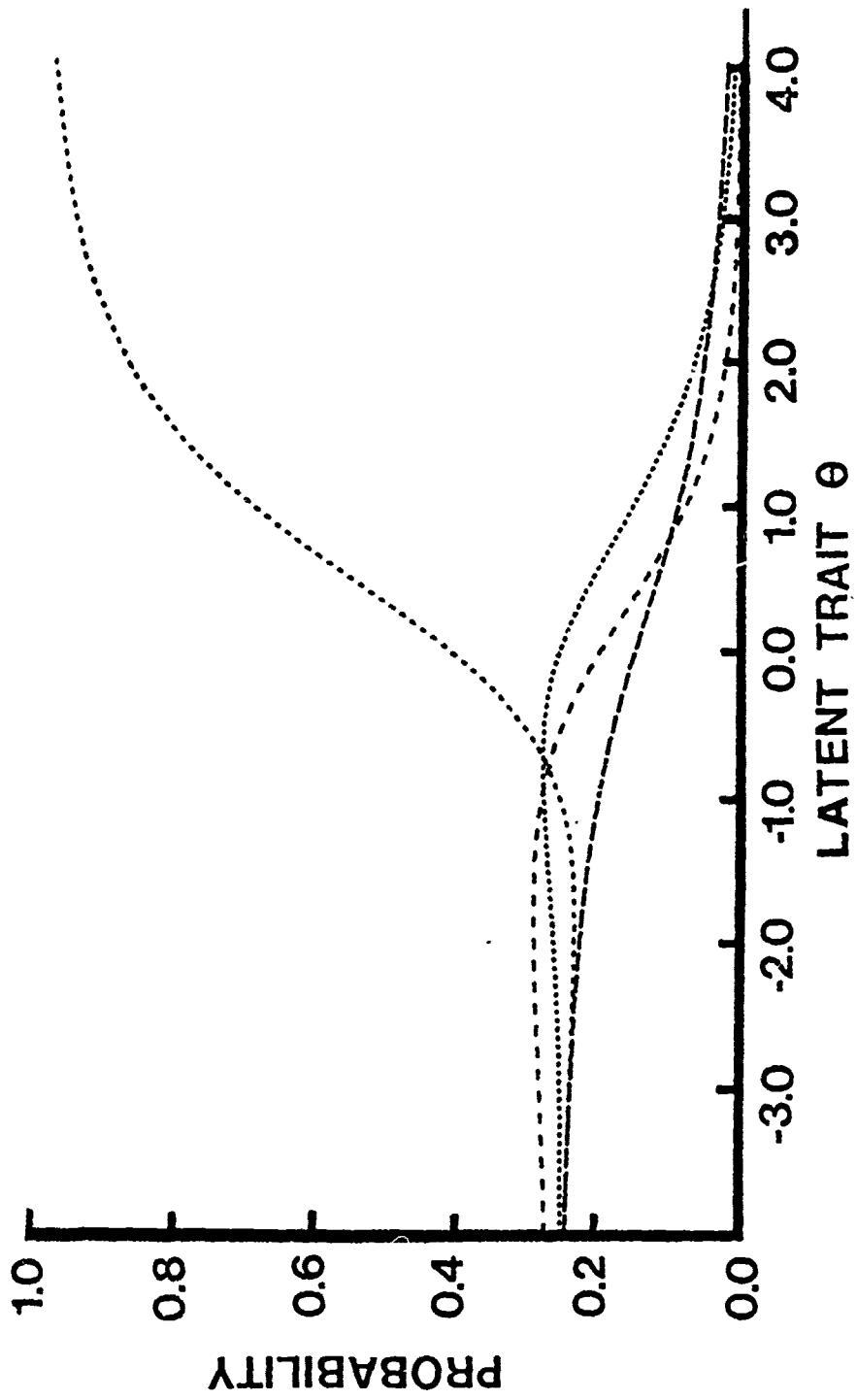


FIGURE 3-13 (Continued)

Operating Characteristics of Four Alternatives Following Model C of Type I  
for the Multiple-Choice Item, with the Same Parameters.

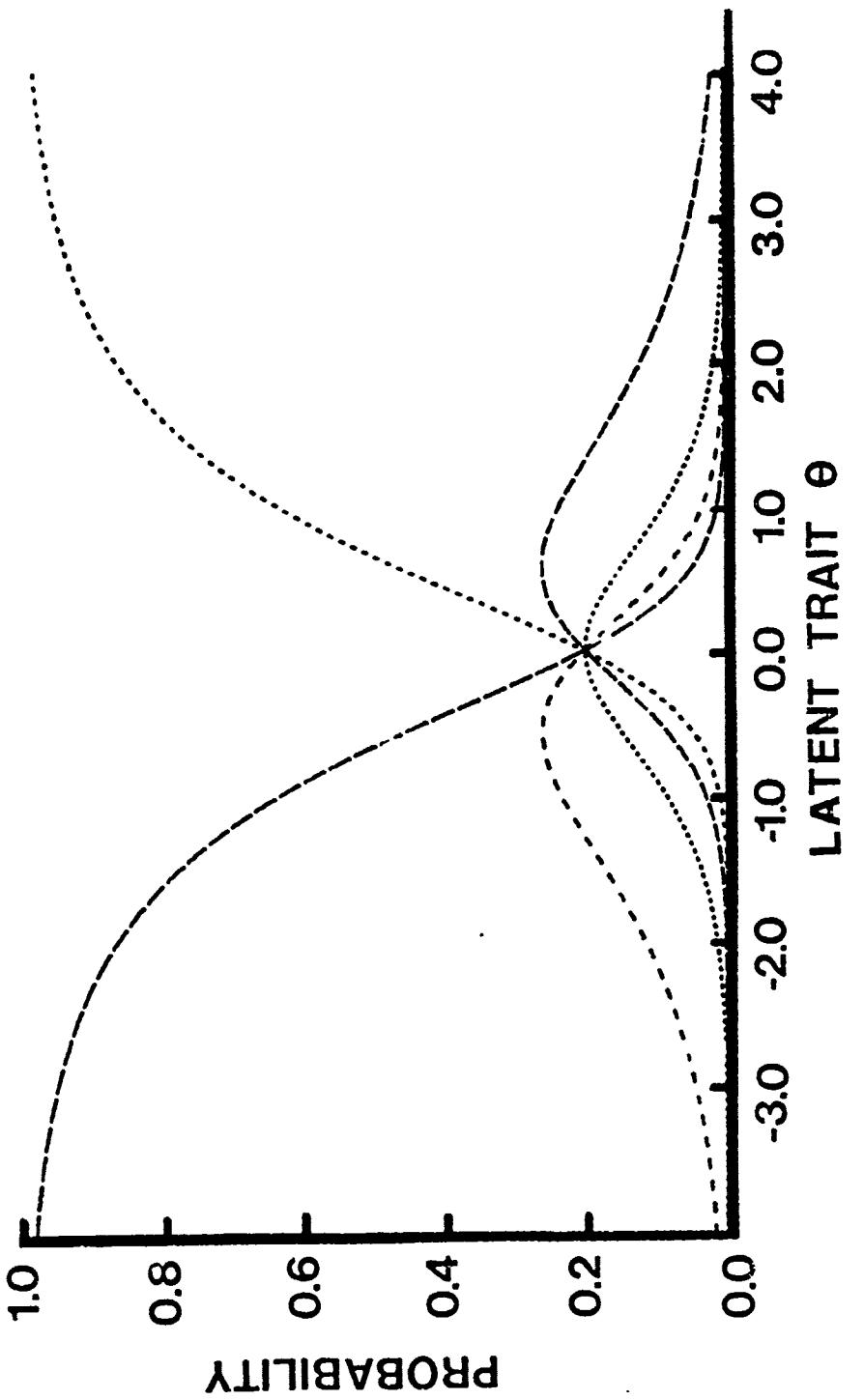


FIGURE 3-14

Operating Characteristics of Five Item Response Categories Following the Multinomial Model. The Item Parameters Are:  $a_1 = -2.000$ ,  $a_2 = -1.000$ ,  $a_3 = 0.000$ ,  $a_4 = 1.000$ ,  $a_5 = 2.000$ ;  $c_1 = 0.000$ ,  $c_2 = 0.000$ ,  $c_3 = 0.000$ ,  $c_4 = 0.000$ ,  $c_5 = 0.000$ .

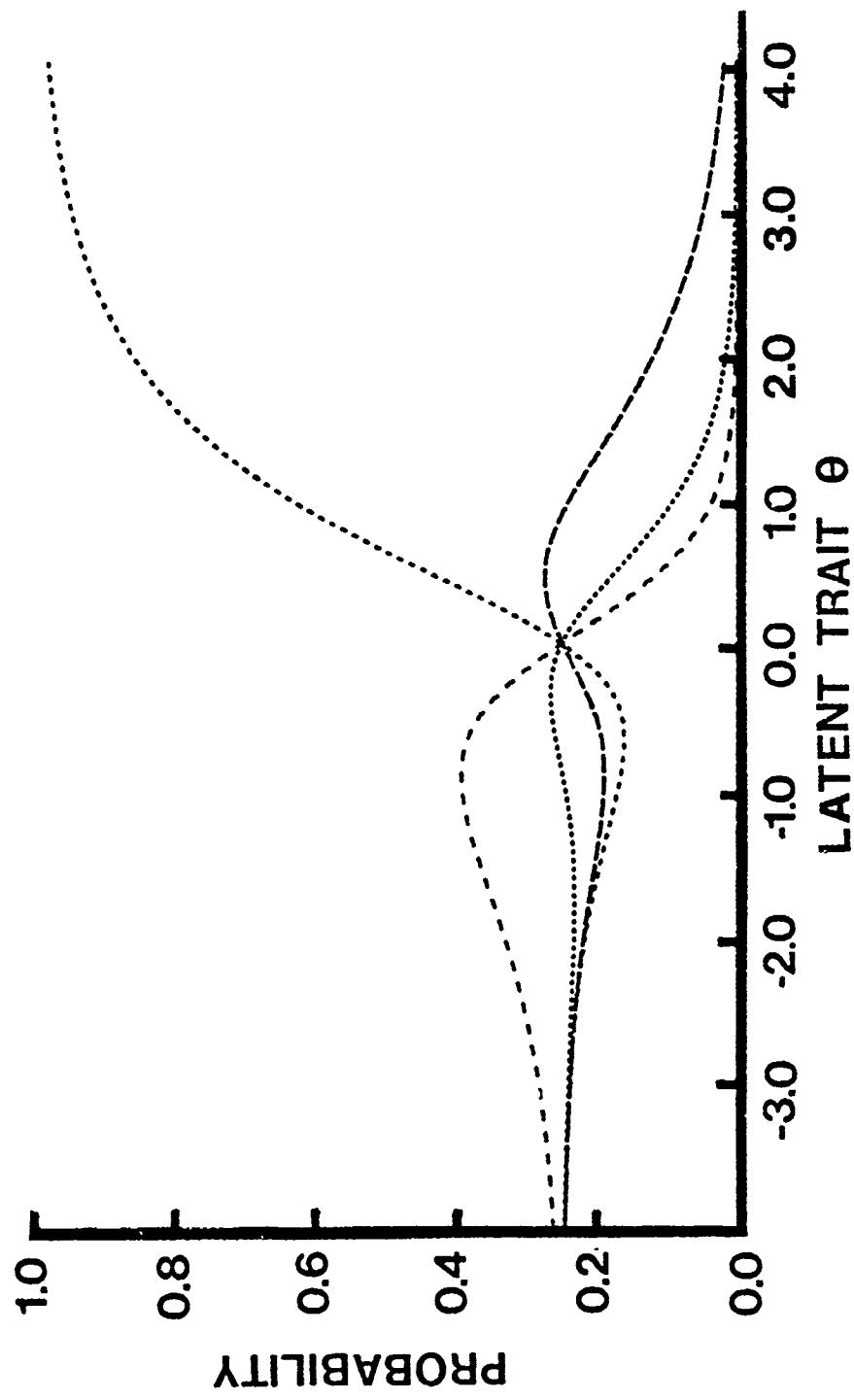


FIGURE 3-14 (Continued)

Operating Characteristics of Four Alternatives Following Model C of Type I  
for the Multiple-Choice Item, with the Same Parameters.

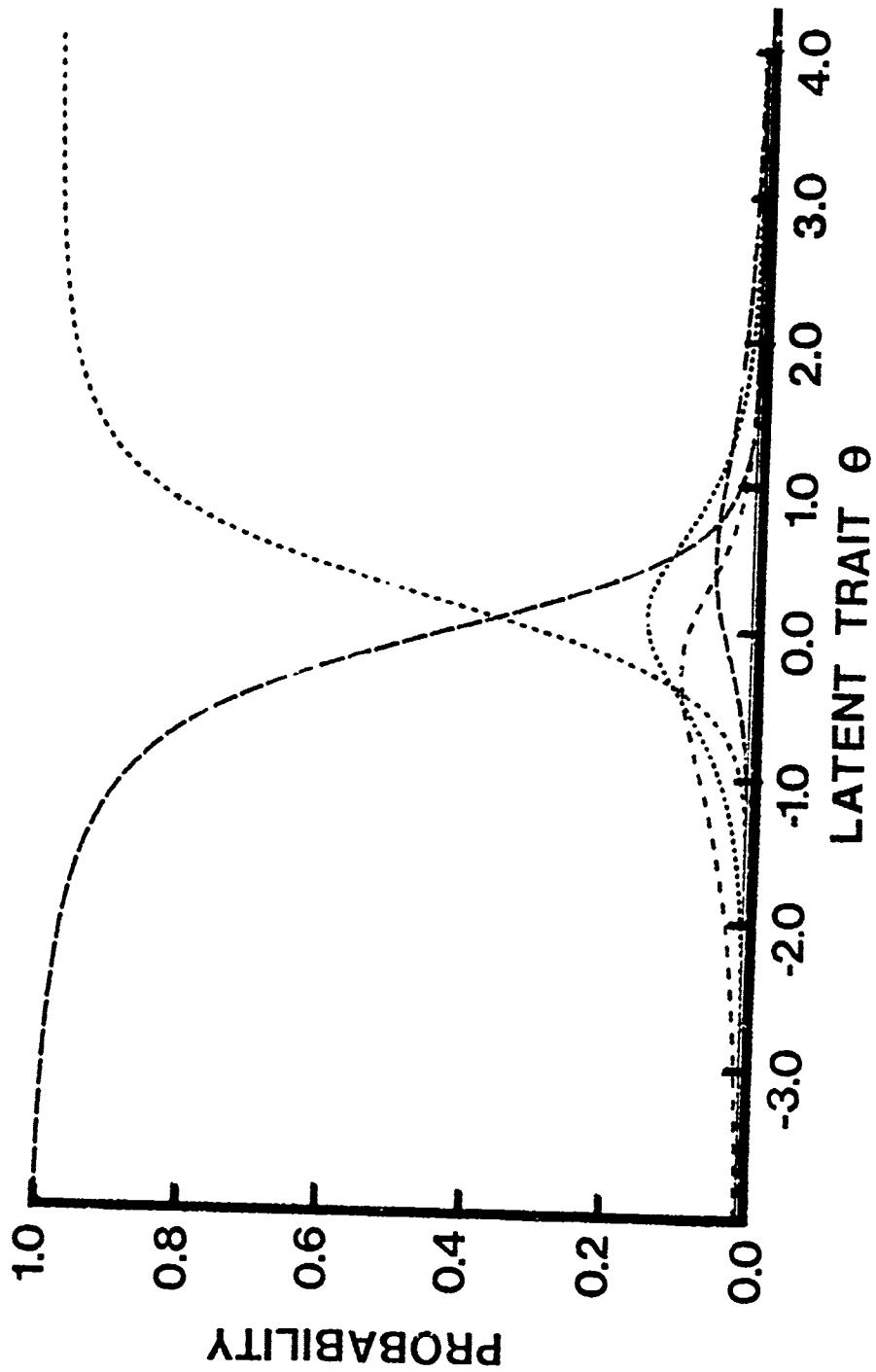


FIGURE 3-15

Operating Characteristics of Five Item Response Categories Following the Multinomial Model. The Item Parameters Are:  $a_1 = -2.000$ ,  $a_2 = -1.000$ ,  $a_3 = 0.000$ ,  $a_4 = 1.000$ ,  $a_5 = 2.000$ ;  $c_1 = 1.000$ ,  $c_2 = -0.500$ ,  $c_3 = 0.000$ ,  $c_4 = -1.250$ ,  $c_5 = 0.750$ .

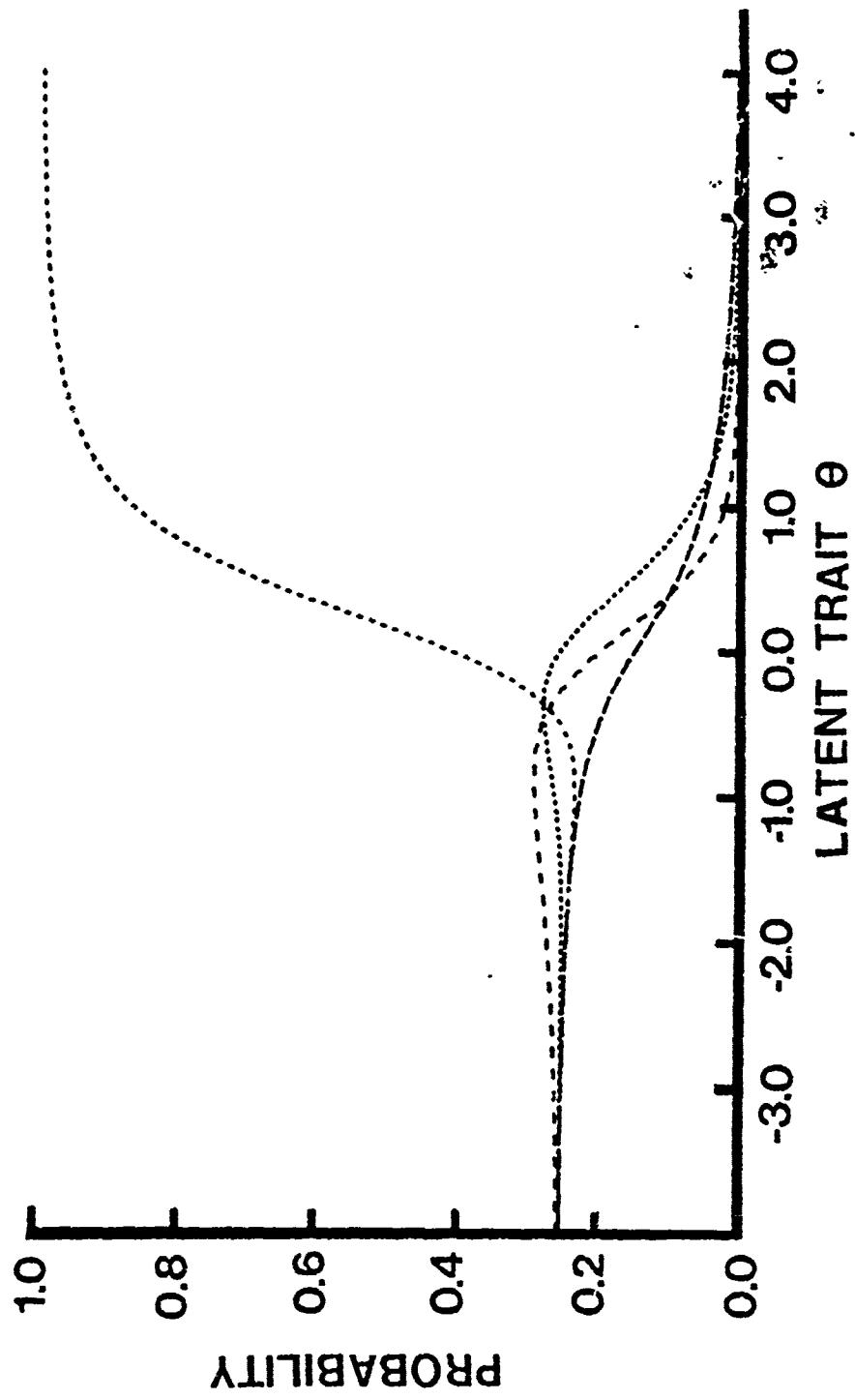
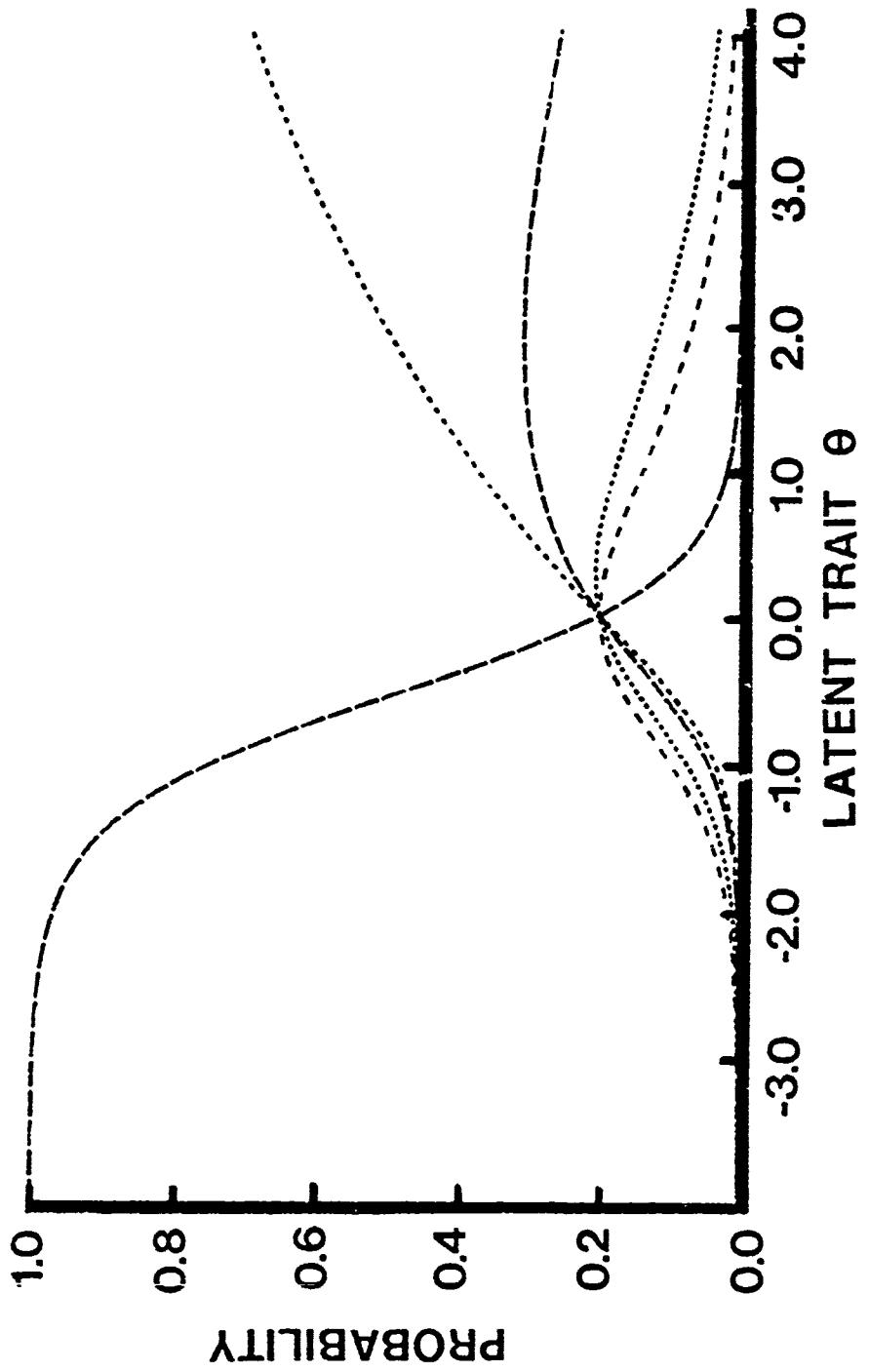


FIGURE 3-15 (Continued)

Operating Characteristics of Four Alternatives Following Model C of Type I  
for the Multiple-Choice Item, with the Same Parameters.



Operating Characteristics of Five Item Response Categories Following the Multinomial Model. The Item Parameters Are:  $a_1 = -2.000, a_2 = 0.000,$   
 $a_3 = 0.250, a_4 = 0.750, a_5 = 1.000 ; c_1 = 0.000, c_2 = 0.000,$   
 $c_3 = 0.000, c_4 = 0.000, c_5 = 0.000 .$

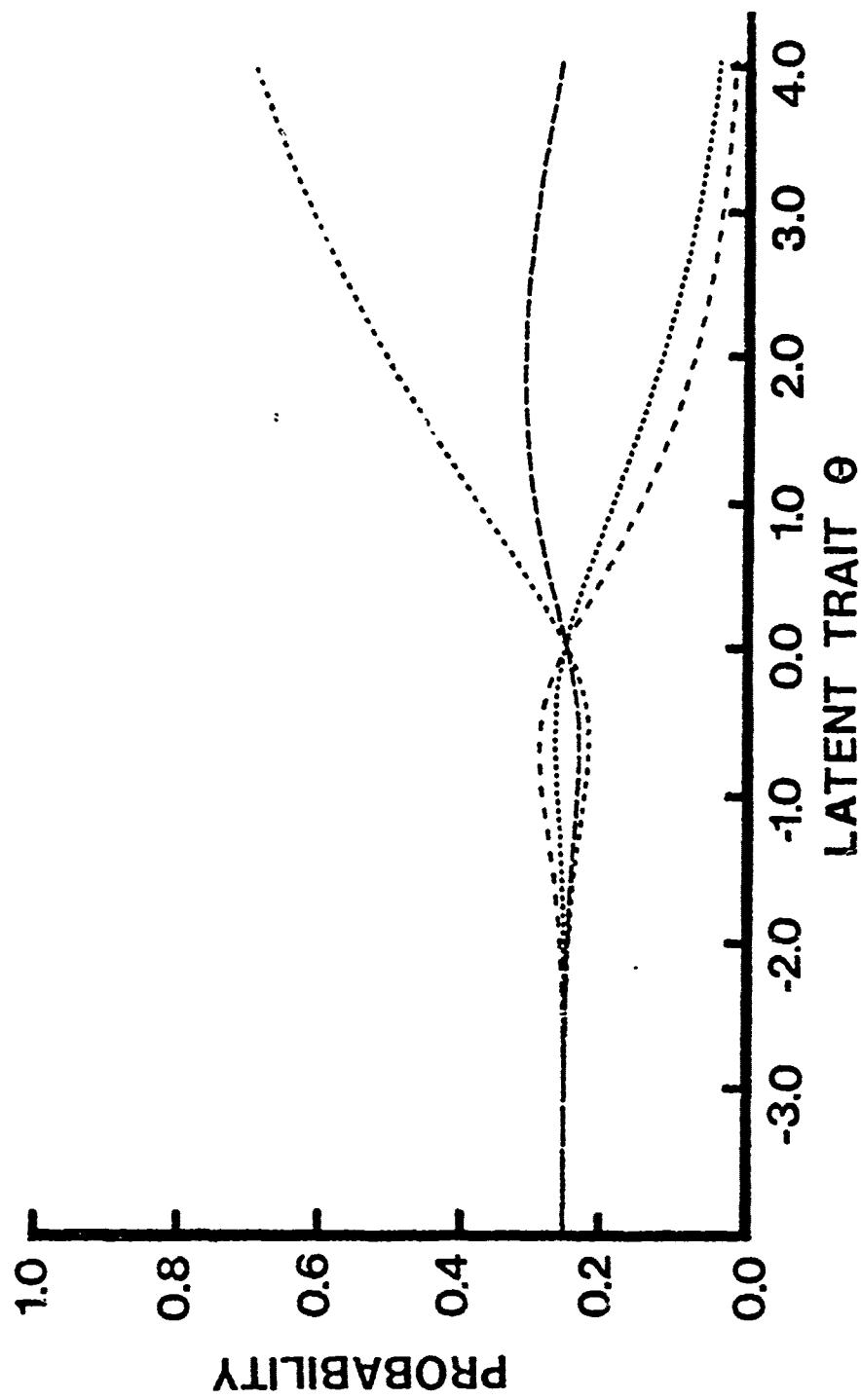


FIGURE 3-16 (Continued)

Operating Characteristics of Four Alternatives Following Model C of Type I  
for the Multiple-Choice Item, with the Same Parameters.

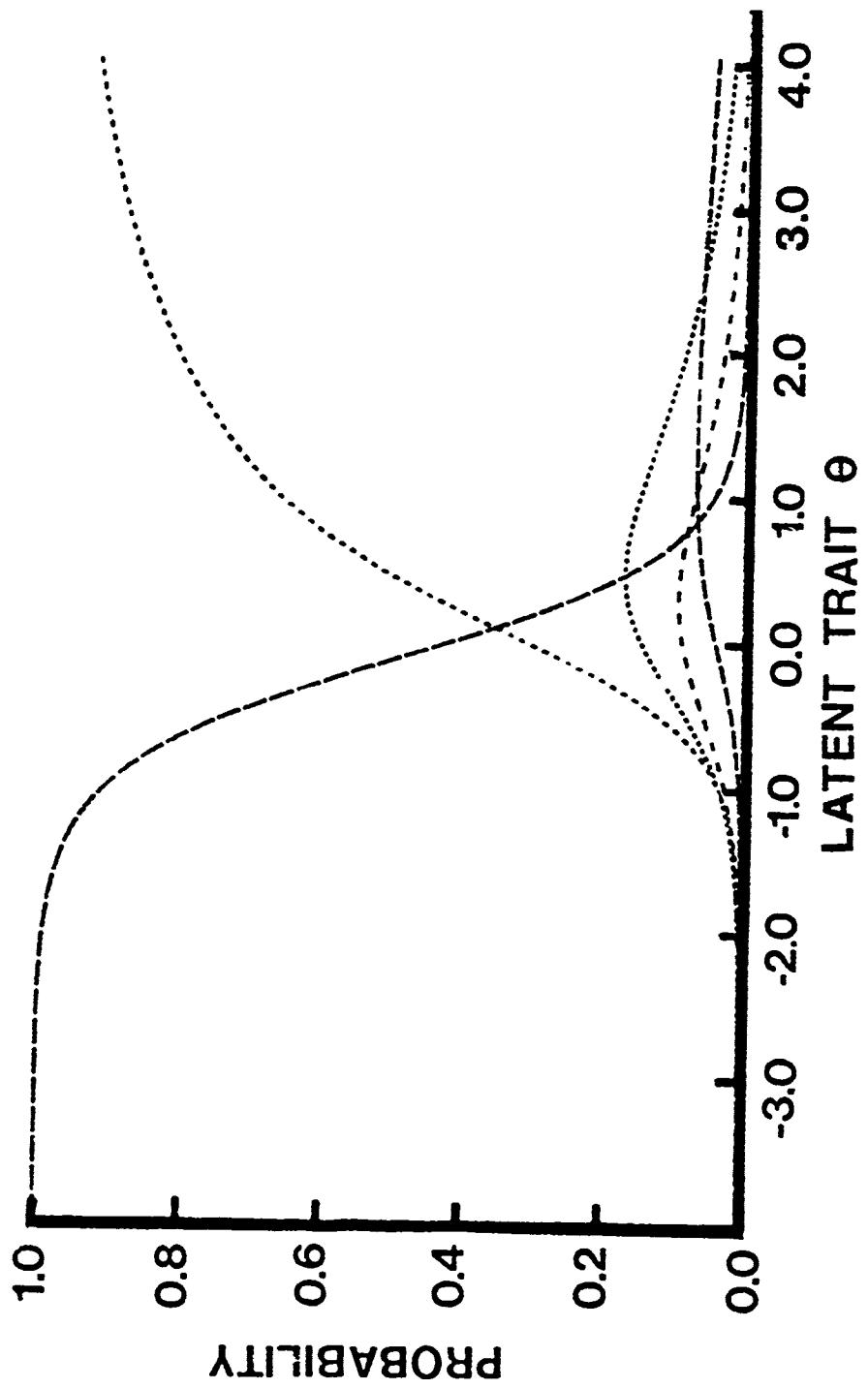


FIGURE 3-17

Operating Characteristics of Five Item Response Categories Following the Multinomial Model. The Item Parameters Are:  $a_1 = -2.000$ ,  $a_2 = 0.000$ ,  $a_3 = 0.250$ ,  $a_4 = 0.750$ ,  $a_5 = 1.000$ ;  $c_1 = 1.000$ ,  $c_2 = -0.500$ ,  $c_3 = 0.000$ ,  $c_4 = -1.250$ ,  $c_5 = 0.750$ .

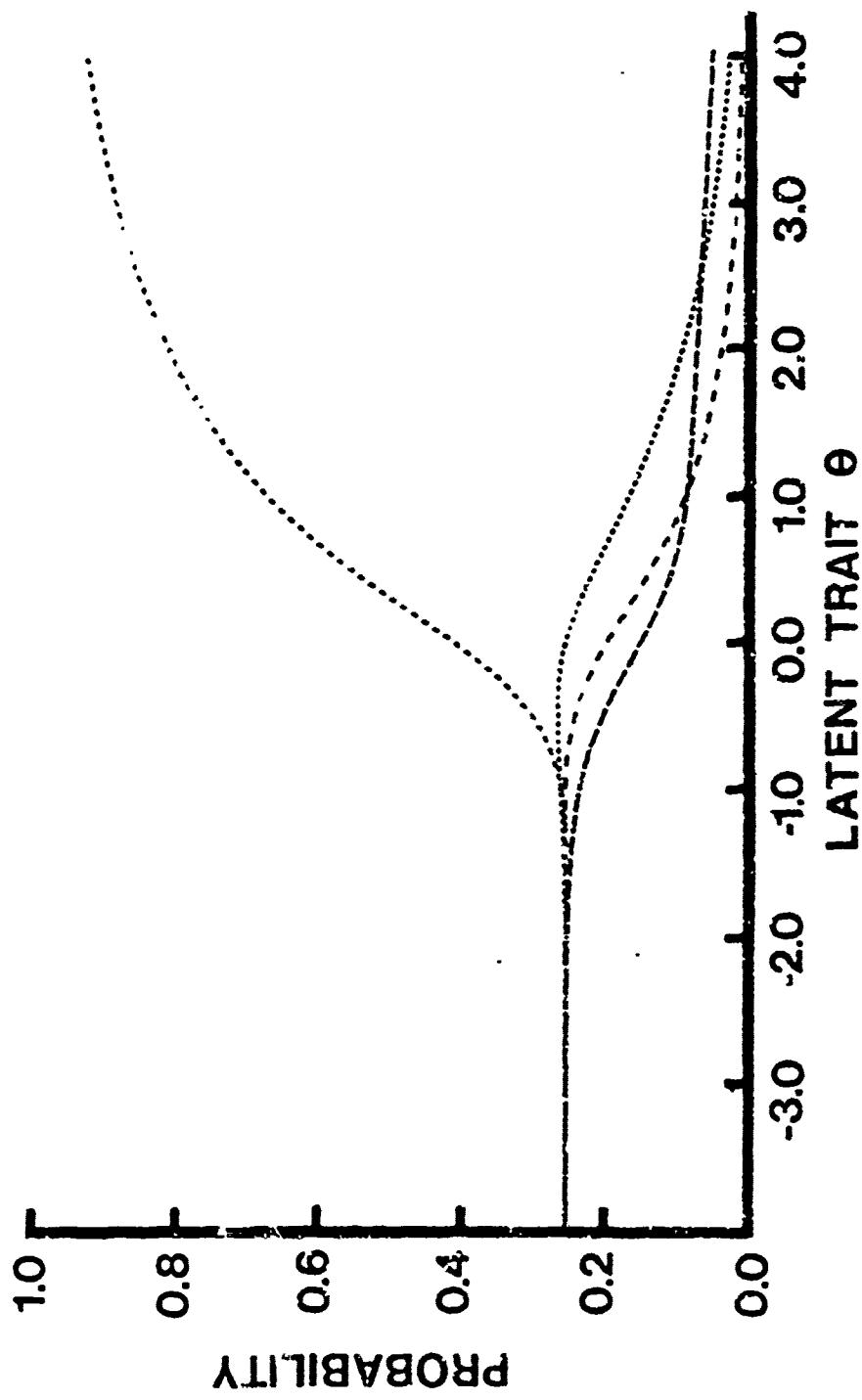


FIGURE 3-17 (Continued)

Operating Characteristics of Four Alternatives Following Model C of Type I  
for the Multiple-Choice Item, with the Same Parameters.

#### IV Basic Functions and Information Functions of the Multiple-Choice Item

The basic function,  $A_{x_g}(\theta)$ , of the item response category  $x_g$ , was defined by Samejima (Samejima, 1969), such that

$$(4.1) \quad A_{x_g}(\theta) = \frac{\partial}{\partial \theta} \log P_{x_g}(\theta) = P'_{x_g}(\theta)/P_{x_g}(\theta),$$

where  $P'_{x_g}(\theta)$  denotes the first derivative of the operating characteristic,  $P_{x_g}(\theta)$ , with respect to ability  $\theta$ . This function has an essential role in the numerical solution of the maximum likelihood estimation of the examinee's ability. A sufficient condition that a model defined on the graded response level provides us with a unique maximum likelihood estimate for every possible response pattern, or unique maximum condition, is that this basic function is strictly decreasing in  $\theta$  with a non-negative asymptote at  $\theta \rightarrow -\infty$  and a non-positive asymptote at  $\theta \rightarrow \infty$ , with respect to every item response category (cf. Samejima, 1969, 1972).

The item response information function,  $I_{x_g}(\theta)$ , of the graded item score  $x_g$  was defined by Samejima (Samejima, 1972) such that

$$(4.2) \quad I_{x_g}(\theta) = - \frac{\partial^2}{\partial \theta^2} \log P_{x_g}(\theta) = - \frac{\partial}{\partial \theta} A_{x_g}(\theta) \\ = [A_{x_g}(\theta)]^2 - [P''_{x_g}(\theta)/P_{x_g}(\theta)],$$

where  $P''_{x_g}(\theta)$  is the second derivative of the operating characteristic  $P_{x_g}(\theta)$  with respect to  $\theta$ . Thus the first half of the unique

maximum condition can be restated in terms of the item response information function, instead of the basic function, i.e., our model provides us with a unique maximum likelihood estimate for every possible response pattern under the model. i: the item response information function is positive, or assumes zero for, at most, an enumerable number of points of  $\theta$  with respect to every item response category, and so forth. It has been shown that both the normal ogive and the logistic models satisfy the unique maximum condition, and so does Bock's multinomial model, whereas this is not the case for the family of three-parameter models (cf. Samejima, 1972, 1973b).

The item information function,  $I_g(\theta)$ , of item  $g$  is defined as the conditional expectation of the item response information function, such that

$$(4.3) \quad I_g(\theta) = \sum_{x=0}^{m_g} I_x(\theta) P_{xg}(\theta) .$$

Since from (4.2) and (4.3) we obtain

$$(4.4) \quad I_g(\theta) = \sum_{x=0}^{m_g} [A_x(\theta)]^2 P_{xg}(\theta) ,$$

we can see that the item information function is always non-negative for the entire range of  $\theta$ , whether the model satisfies the unique maximum condition, or not. Thus the use of the item information function, or the test information function, as a measure of the accuracy of ability estimation when the three-parameter logistic model is adopted is meaningless and deceptive, as was pointed out

earlier (Samejima, 1973b).

The analogous basic function and information functions can be defined for the multiple-choice item. Thus the basic function,  $A_h(\theta)$ , of the alternative  $h$  is written as

$$(4.5) \quad A_h(\theta) = \frac{\partial}{\partial \theta} \log P_h(\theta) = P'_h(\theta)/P_h(\theta) ,$$

where  $P'_h(\theta)$  is the first derivative of the operating characteristic  $P_h(\theta)$  of the alternative  $h$  with respect to ability  $\theta$ . For the alternative information function,  $I_h(\theta)$ , we can write

$$(4.6) \quad I_h(\theta) = - \frac{\partial^2}{\partial \theta^2} \log P_h(\theta) = [A_h(\theta)]^2 - P''_h(\theta)/P_h(\theta) ,$$

where  $P''_h(\theta)$  is the second derivative of the operating characteristic  $P_h(\theta)$  of the alternative  $h$  with respect to  $\theta$ . The item information function of the multiple-choice item is the conditional expectation of the alternative information function, given  $\theta$ , such that

$$(4.7) \quad I_g(\theta) = \sum_{h=1}^m I_h(\theta) P_h(\theta) = \sum_{h=1}^m [A_h(\theta)]^2 P_h(\theta) .$$

It should be noted that these basic functions and information functions assume more complicated forms than the corresponding functions for the graded item response categories, if we adopt one of the models for the multiple-choice item proposed in Chapter 3. Because of the popularity of the three-parameter logistic model among researchers, we shall take Model B of Type I in this chapter, and observe its basic functions and information functions, which are given by (4.5), (4.6) and (4.7), with respect to the

examples given in Chapter 3 . Comparison will be made between these functions in Model B and those in the logistic model on the graded response level, which share the same parameters, as we did in the preceding chapter.

Let  $P_h^*(\theta)$  be such that

$$(4.8) \quad P_h^*(\theta) = [1 + \exp\{-Da(\theta - b_h)\}]^{-1} .$$

Then we can rewrite (3.4) for the operating characteristic of the alternative  $h$  in the form

$$(4.9) \quad P_h(\theta) = [1 - \exp\{-Da(b_{h+1} - b_h)\}]P_h^*(\theta)[1 - P_{h+1}^*(\theta)] \\ + (1/m)[1 - P_1^*(\theta)] .$$

From (4.9) we obtain the first and second derivatives of  $P_h(\theta)$  such that

$$(4.10) \quad P'_h(\theta) = Da[\{1 - \exp\{-Da(b_{h+1} - b_h)\}\}P_h^*(\theta)\{1 - P_{h+1}^*(\theta)\}\{1 - P_h^*(\theta) - P_{h+1}^*(\theta)\}] \\ - (1/m)DaP_1^*(\theta)[1 - P_1^*(\theta)]$$

and

$$(4.11) \quad P''_h(\theta) = D^2a^2[\{1 - \exp\{-Da(b_{h+1} - b_h)\}\}\{1 - P_h^*(\theta) - P_{h+1}^*(\theta)\}^2 \\ - P_h^*(\theta)[1 - P_h^*(\theta)] - P_{h+1}^*(\theta)[1 - P_{h+1}^*(\theta)\}] \\ - (1/m)D^2a^2P_1^*(\theta)[1 - P_1^*(\theta)][1 - 2P_1^*(\theta)] .$$

It is noted that the last term in each of (4.9), (4.10) and (4.11) is the term which makes the function different from the corresponding function in the logistic model on the graded response level. The

amount of effect caused by these additional terms on the basic functions and the information functions for different levels of  $\theta$  depends upon the parameter  $b_1$ , or  $b_h$  for  $h = 1$ . If these additional terms do not exist, i.e., in the logistic model on the graded response level, we can write for the basic functions and the information functions

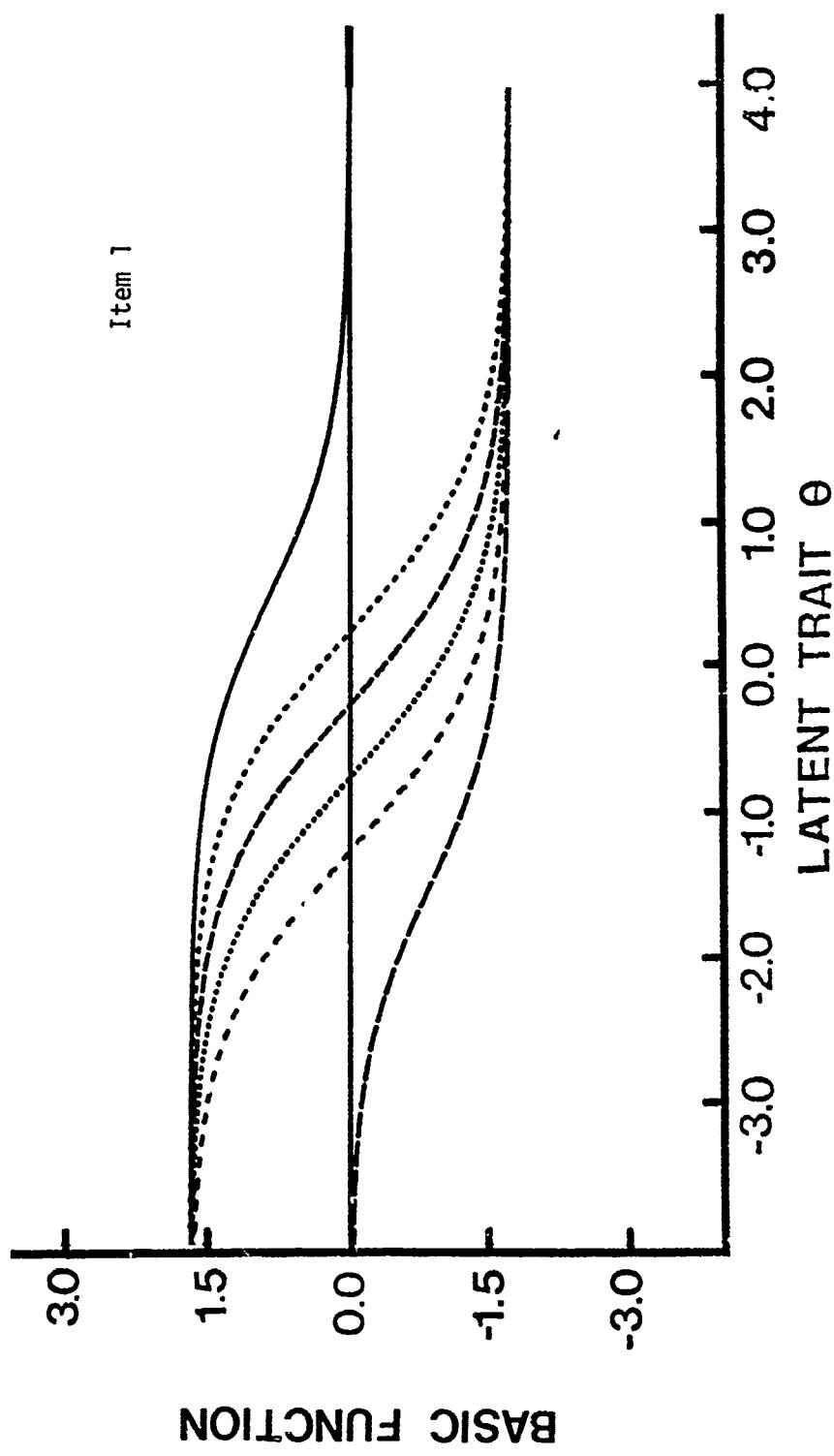
$$(4.12) \quad A_h(\theta) = Da[1-P_h^*(\theta)-P_{h+1}^*(\theta)] ,$$

$$(4.13) \quad I_h(\theta) = D^2a^2[P_h^*(\theta)\{1-P_h^*(\theta)\} + P_{h+1}^*(\theta)\{1-P_{h+1}^*(\theta)\}] ,$$

where  $h = 0, 1, 2, \dots, m$ , and

$$(4.14) \quad I_g(\theta) = D^2a^2 \sum_{h=0}^m [1-P_h^*(\theta)-P_{h+1}^*(\theta)]^2 [P_h^*(\theta)-P_{h+1}^*(\theta)] .$$

Figures 4-1 through 4-5 present the basic functions of the six categories in the logistic model, which are given by (4.12), and those of the five alternatives in Model B of Type I for the multiple-choice item, which were obtained by substituting (4.9) and (4.10) into (4.5), for the five hypothetical test items observed in the preceding chapter. As we can see in the first graph of each figure, all the six basic functions in the logistic model are strictly decreasing in  $\theta$ , with the common asymptote,  $1.7a$ , at  $\theta \rightarrow -\infty$  for  $h = 2, 3, 4, 5, 6$  and  $-1.7a$  at  $\theta \rightarrow \infty$  for  $h = 1, 2, 3, 4, 5$ , while for  $h = 6$  the asymptote at  $\theta \rightarrow \infty$  is zero and for  $h = 1$  the one at  $\theta \rightarrow -\infty$  is zero, respectively (cf. Samejima, 1969). It should also be noted that for the four intermediate categories,  $h = 2, 3, 4, 5$ , the basic functions take on zero at  $\theta = (b_h + b_{h+1})/2$ .



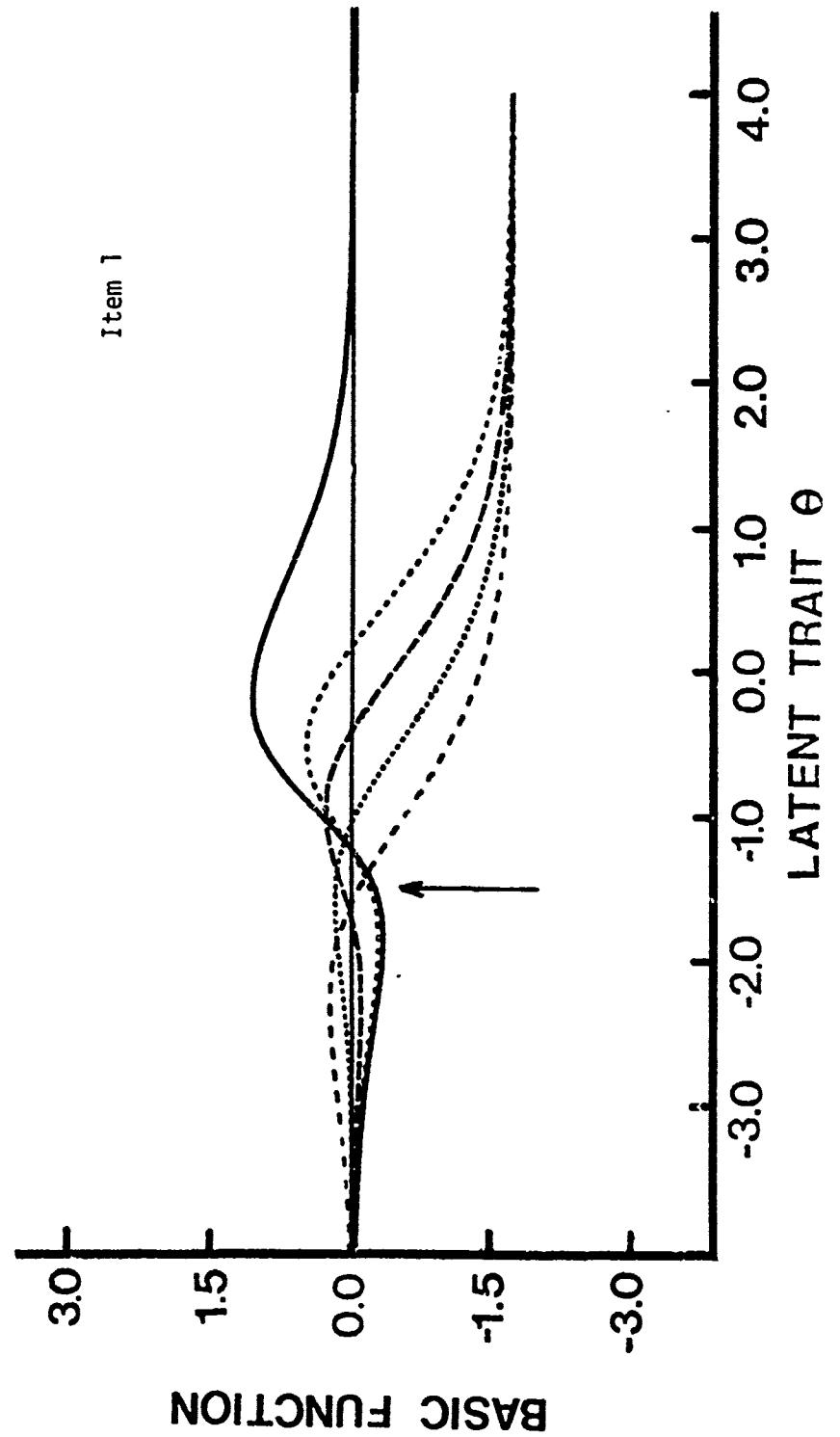
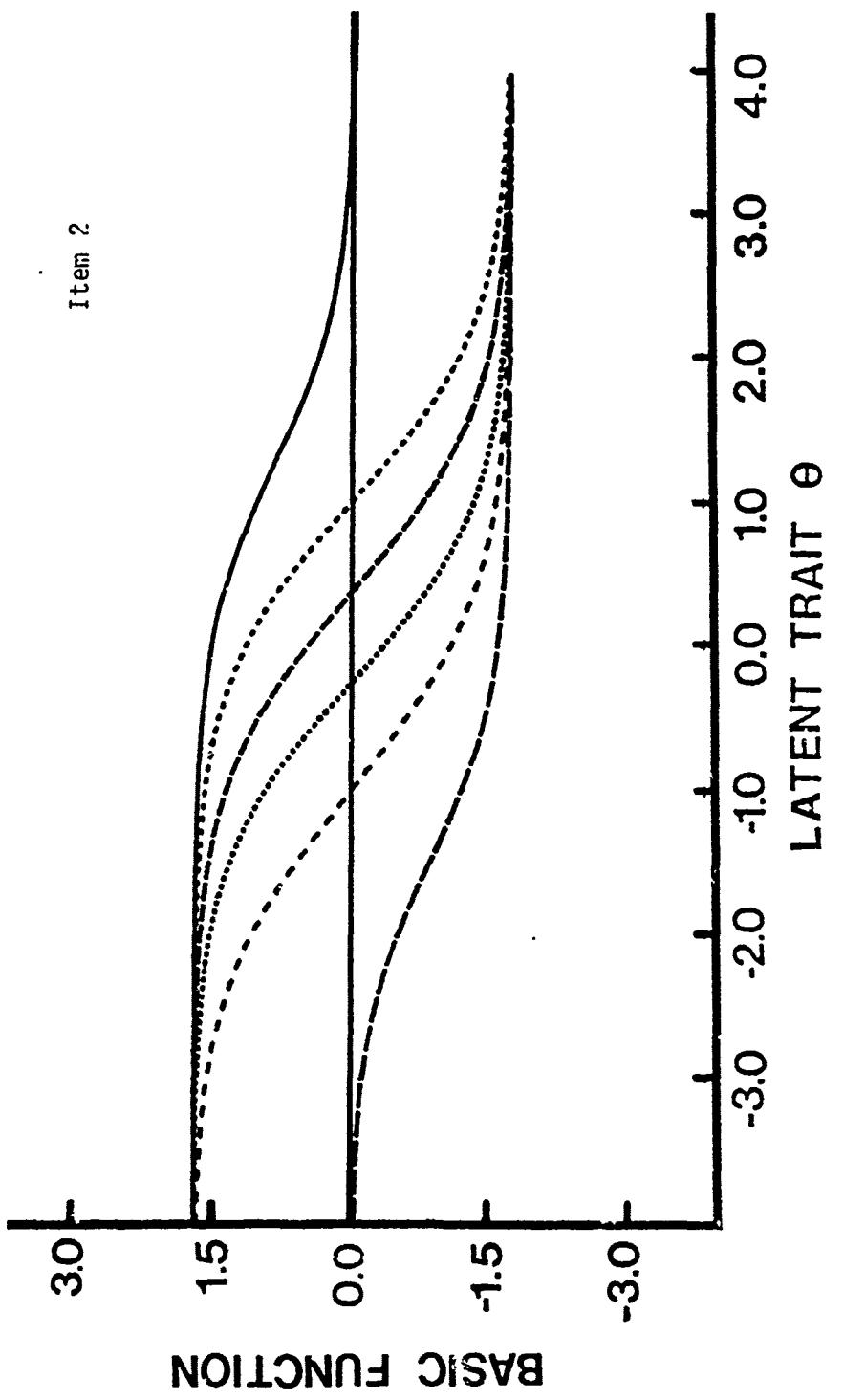


FIGURE 4-1

Basic Functions of Six Item Response Categories in the Logistic Model (Above), and Those of Five Alternatives in Model B (Below). The Item Parameters Are:  $a_g = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -1.00$ ,  $b_3 = -0.50$ ,  $b_4 = 0.00$  and  $b_5 = 0.50$ .



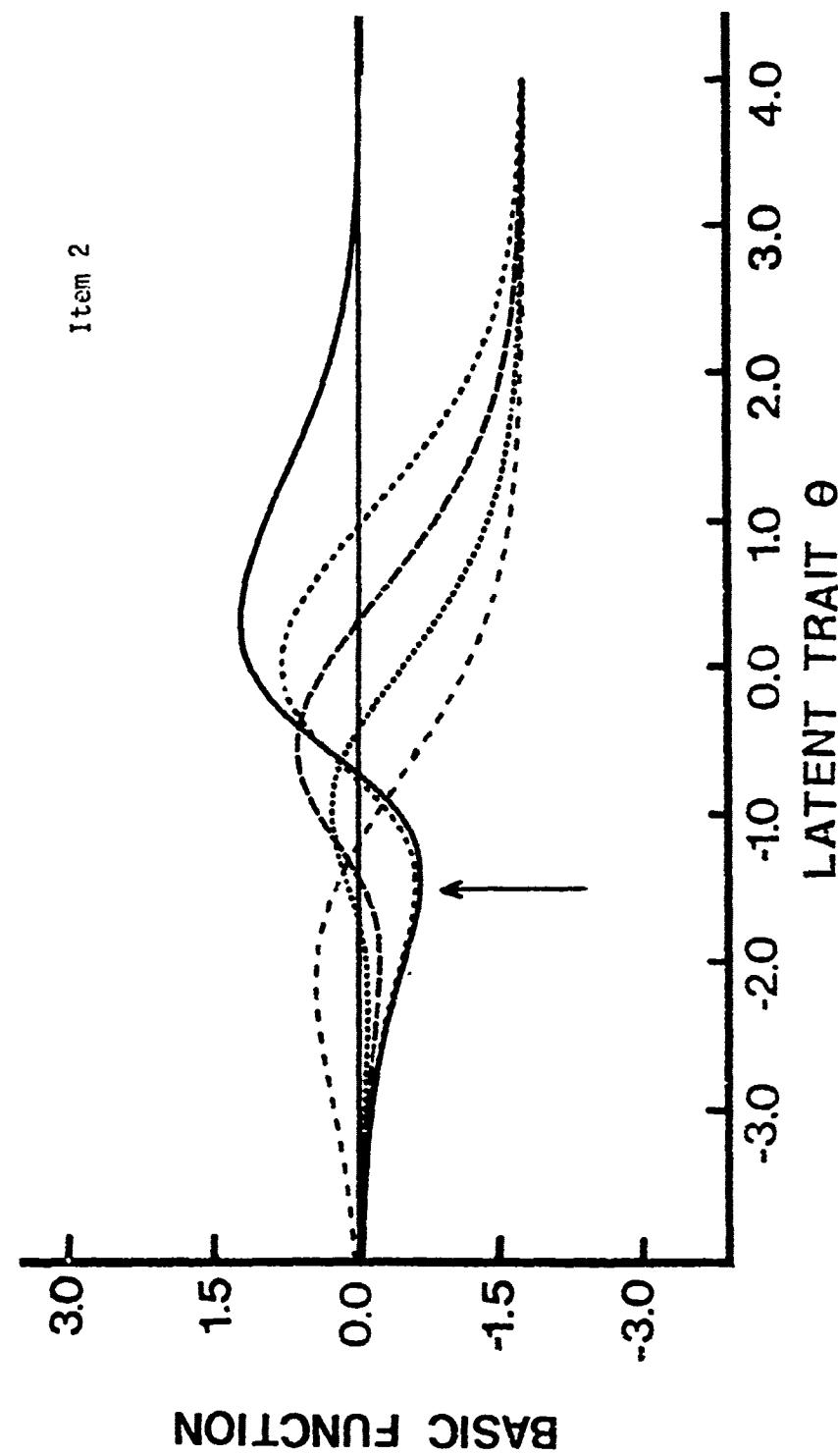
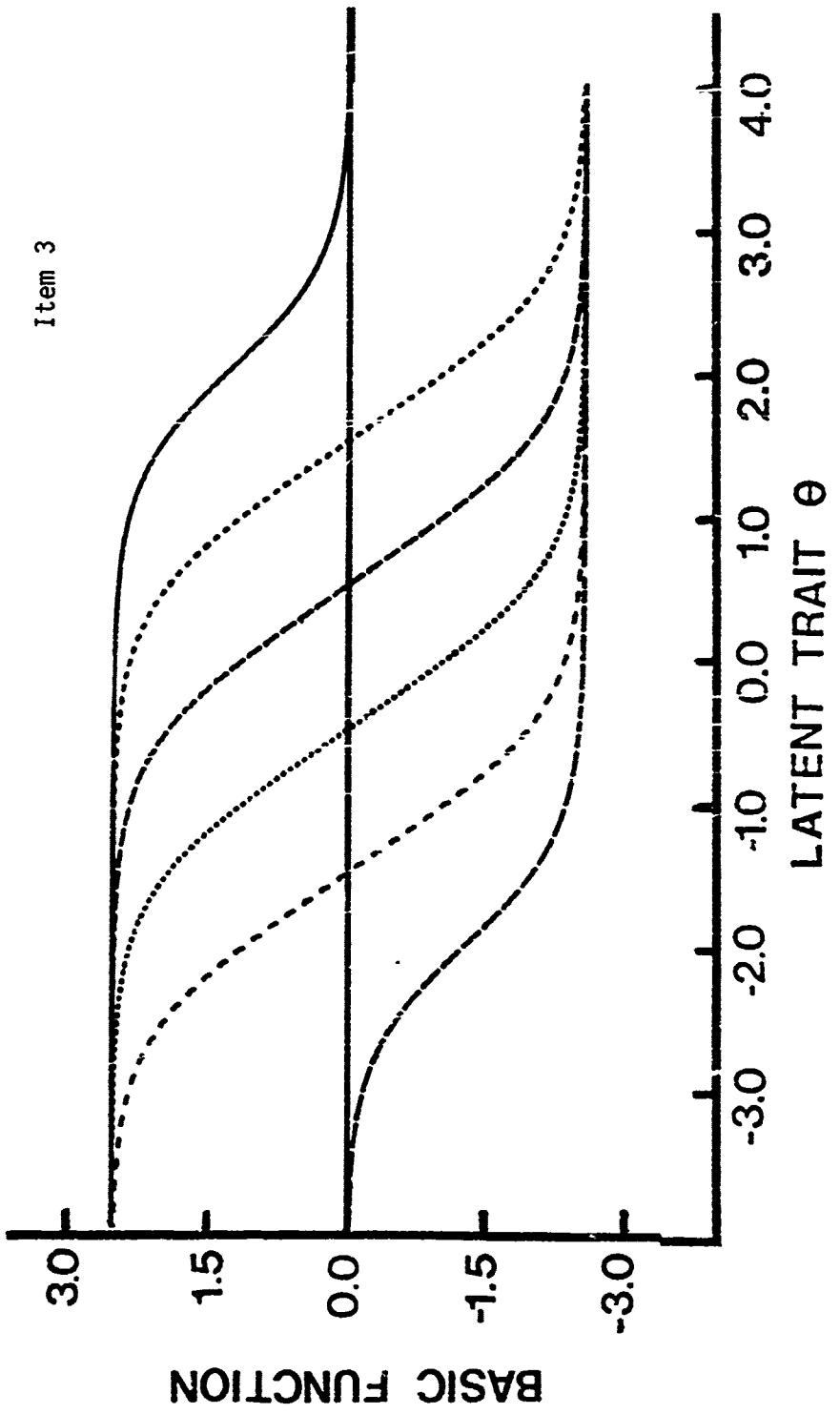


FIGURE 4-2

Basic Functions of Six Item Response Categories in the Logistic Model (Above), and Those of Five Alternatives in Model B (Below). The Item Parameters Are:  $a_g = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -0.50$ ,  $b_3 = 0.00$ ,  $b_4 = 0.75$  and  $b_5 = 1.25$ .



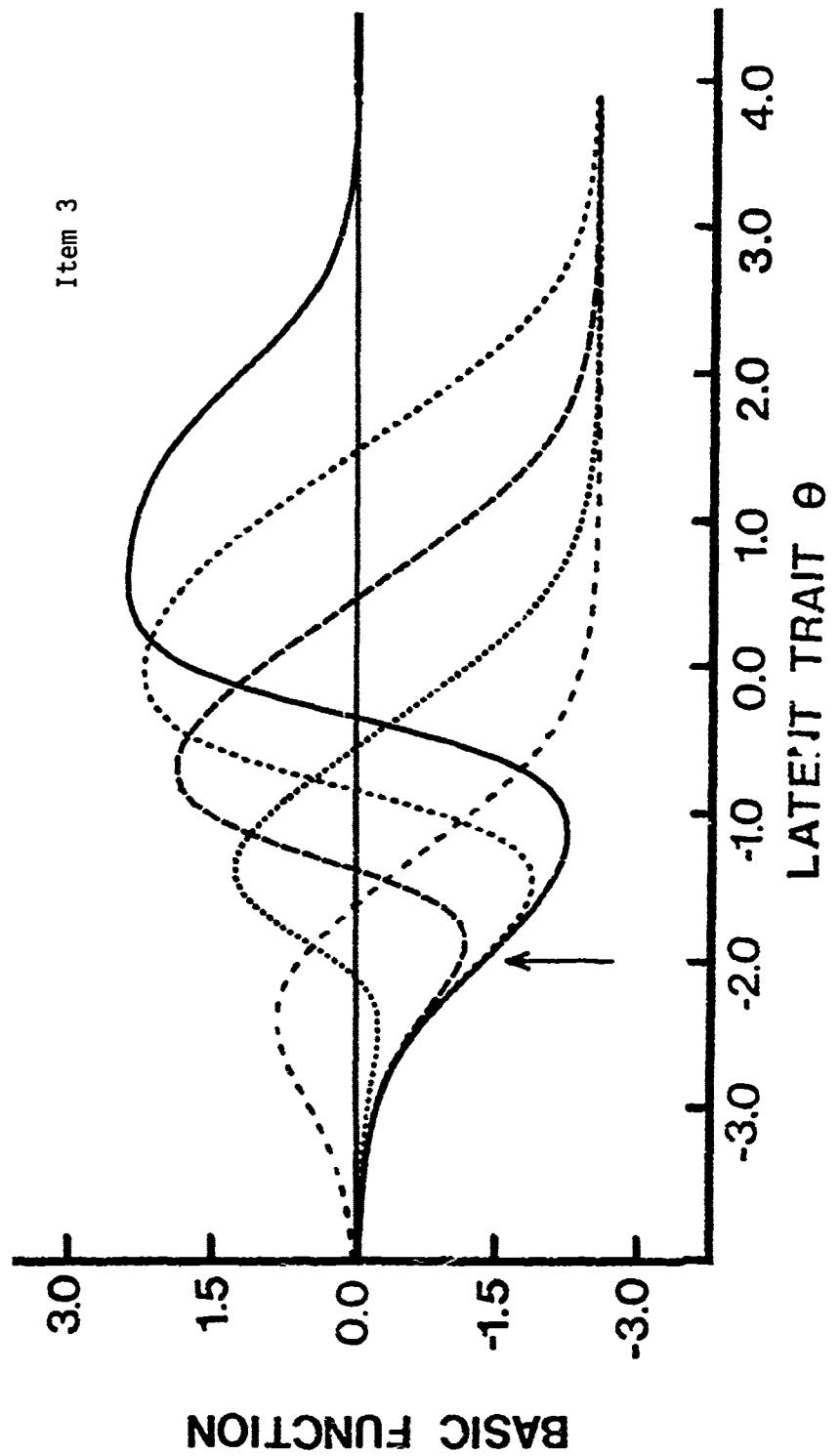
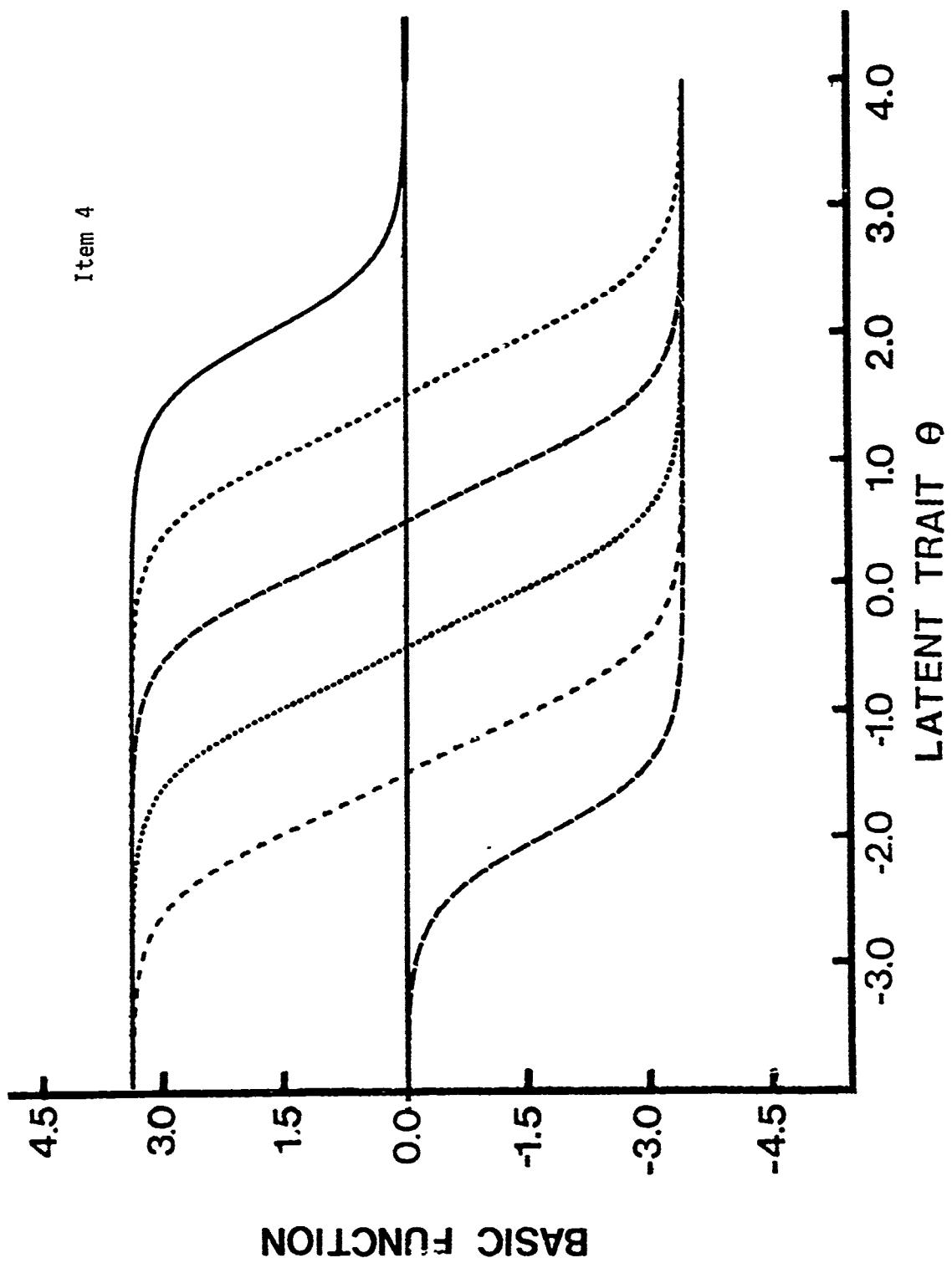


FIGURE 4-3

Basic Functions of Six Item Response Categories in the Logistic Model (Above), and Those of Five Alternatives in Model B (Below). The Item Parameters Are:  $a_g = 1.50$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .



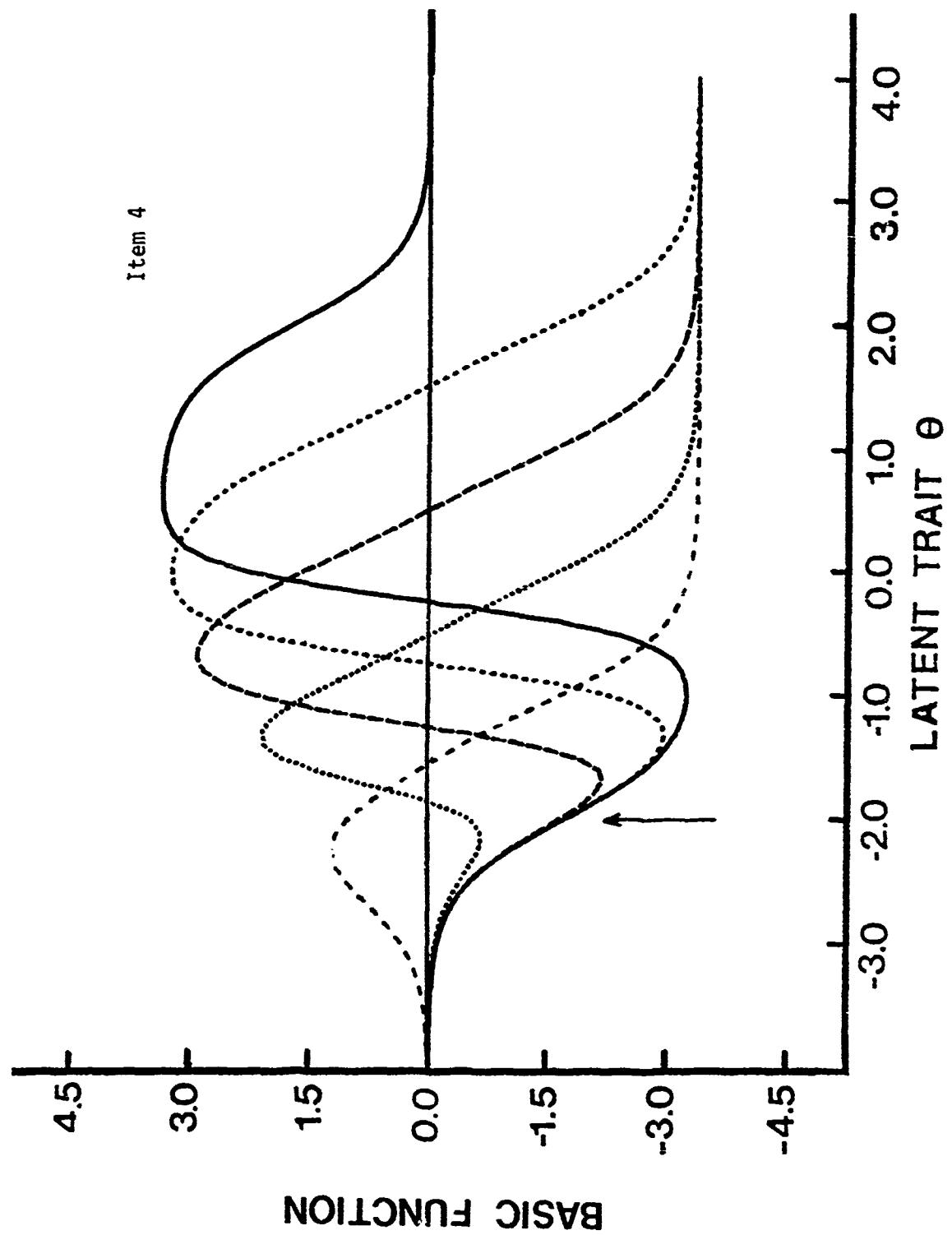
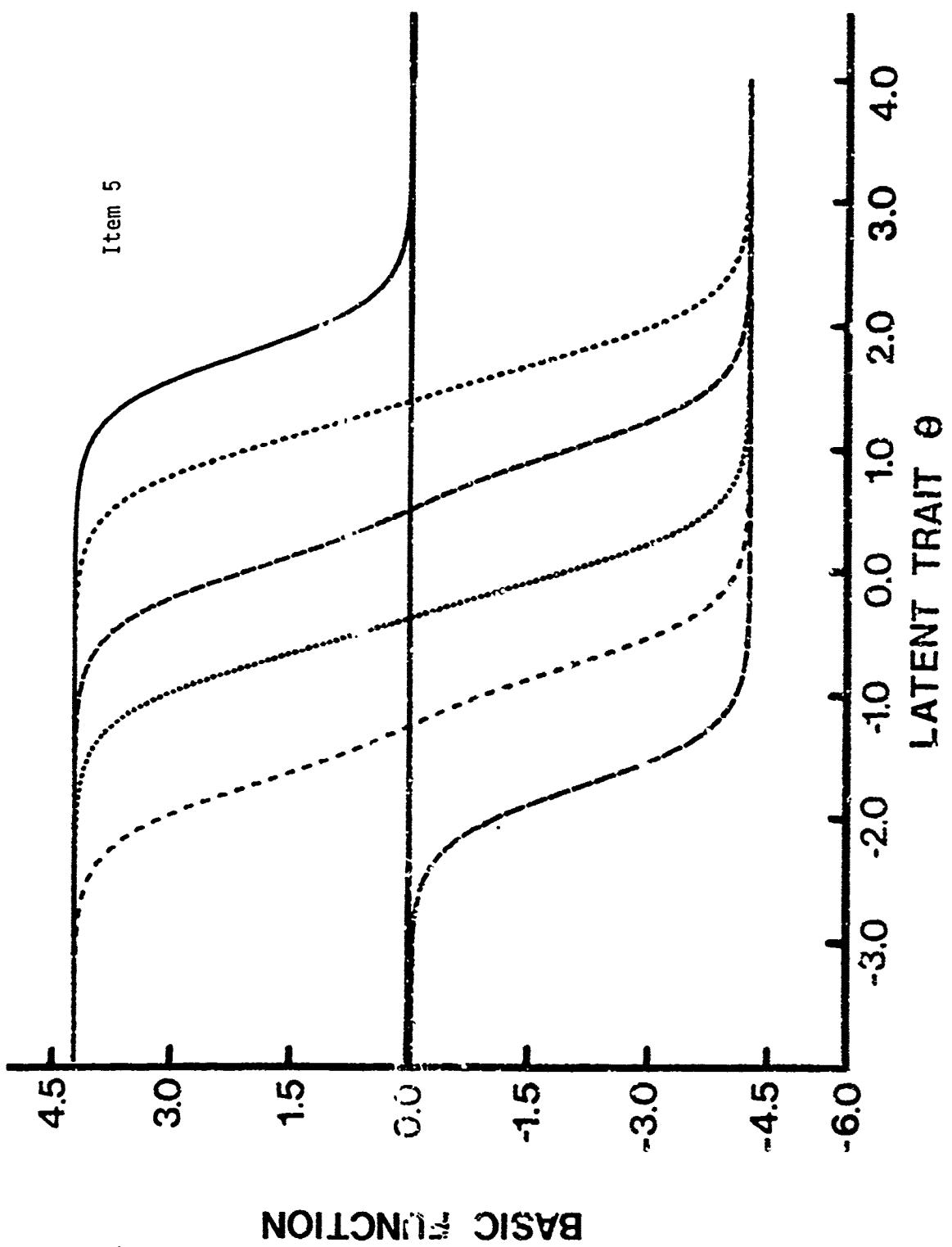


FIGURE 4-4

Basic Functions of Six Item Response Categories in the Logistic Model (Above), and Those of Five Alternatives in Model B (Below). The Item Parameters Are:  $a_g = 2.00$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .



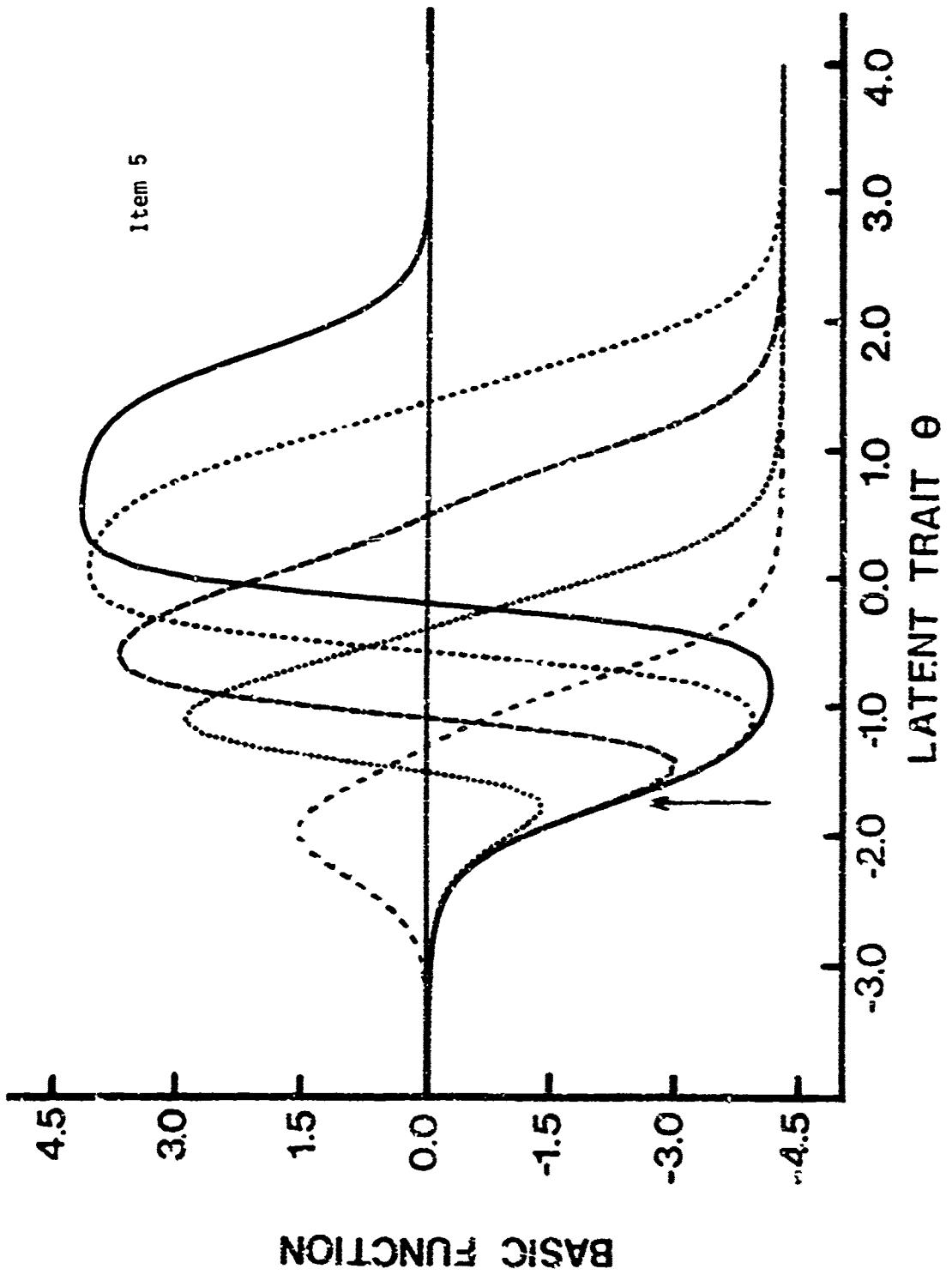


FIGURE 4-5

Basic Functions of Six Item Response Categories in the Logistic Model (Above), and Those of Five Alternatives in Model B (Below). The Item Parameters Are:  $a_8 = 2.50$ ,  $b_1 = -1.75$ ,  $b_2 = -0.75$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 1.75$ .

We find quite a contrasting set of five basic functions in the second graph of each of the five figures, Figures 4-1 through 4-5. In fact, none of these basic functions are strictly decreasing in  $\theta$ , but each has a unique modal point, and, except for  $n = 1$  a unique local minimum also. The common asymptote at  $\theta \rightarrow -\infty$  for the alternatives excluding the correct answer is -1.7a, just as in the logistic model. and the other common asymptote at  $\theta \rightarrow +\infty$ , along with the asymptote at  $\theta \rightarrow -\infty$  for the correct answer, is zero, as is expected from (4.9) and (4.10). It is very obvious from these results that Model B does not satisfy the unique maximum condition, and, therefore, a unique maximum likelihood estimate is not assured for every possible response pattern. We need to pursue the characteristics of this model further and find out some practical solution for this problem, therefore, as was done for the three-parameter logistic model (Samejima, 1973b).

We notice that these basic functions are practically identical with the corresponding curves in the logistic model, for certain intervals of higher ability. Needless to say, it is desirable if these intervals start from relatively lower levels of ability  $\theta$ . It is obvious that the lower endpoint of such an interval depends upon the parameter  $b_1$ , which is indicated by an arrow in each graph of Model B. We can also observe that there is a tendency that this lower endpoint of the interval is higher for an item with a high discrimination parameter.

Figures 4-6 and 4-7 present the alternative information functions in the logistic model, and the corresponding alternative

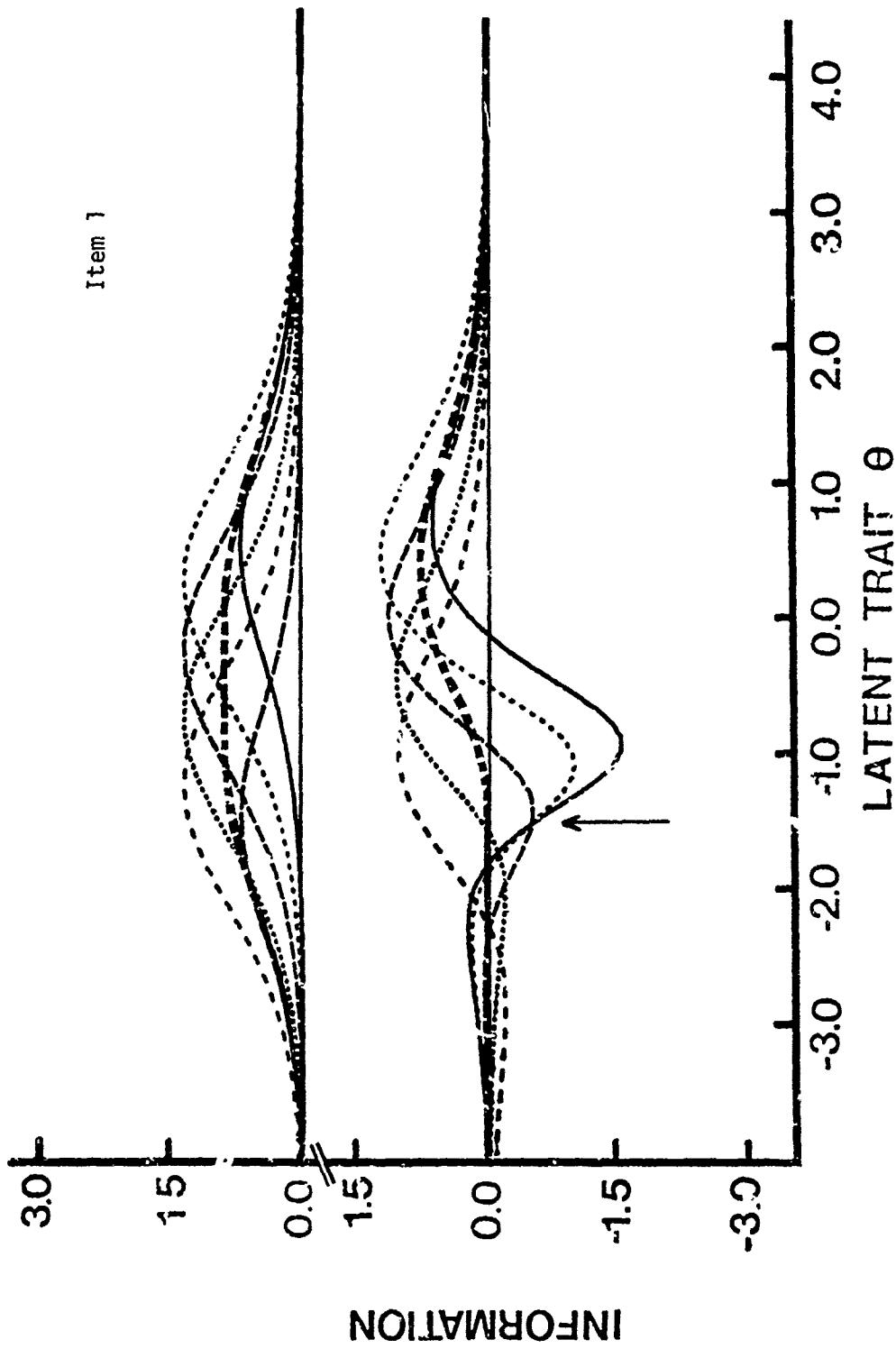


FIGURE 4-6

Item Response Information Functions (Various Thinner Curves) and the Item Information Function (Heavy Dashes) in the Logistic Model (Above) and Those in Model B of Type I for the Multiple-Choice Item (Below). The Item Parameters Are:  $a_g = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -1.00$ ,  $b_3 = -0.50$ ,  $b_4 = \dots$  and  $b_5 = 0.50$ .

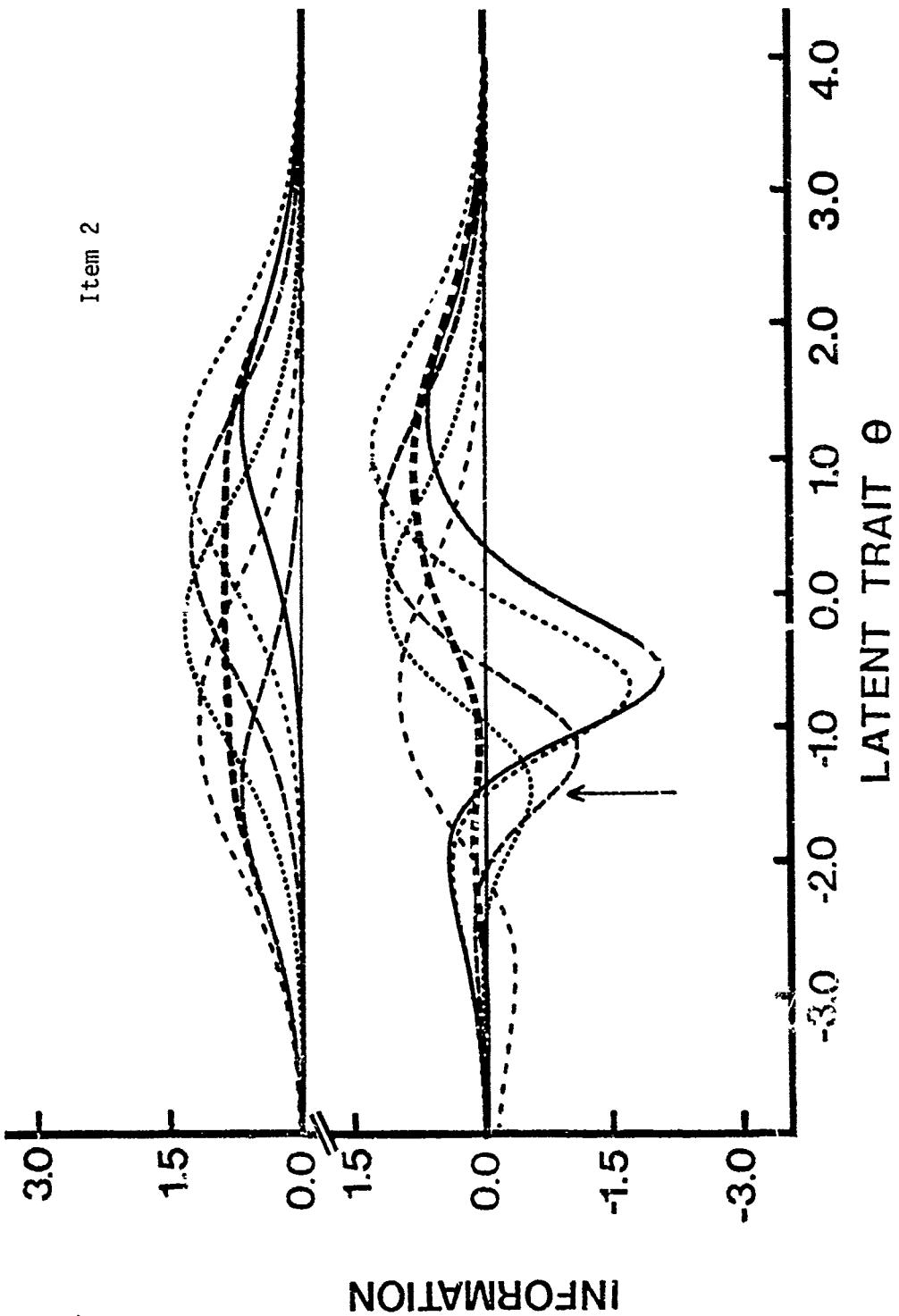


FIGURE 4-7

Item Response Information Functions (Various Thinner Curves) and the Item Information Function (Heavy Dashes) in the Logistic Model (Above) and Those in Model B of Type I for the Multiple-Choice Item (Below). The Item Parameters Are:  $a_g = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -0.50$ ,  $b_3 = 0.00$ ,  $b_4 = 0.75$  and  $b_5 = 1.25$ .

information functions in Model B, in the upper and lower parts, respectively, for items 1 and 2 . Among each of the two sets of six and five curves for the alternative information functions, we find the item information function, which is drawn by a thicker, dashed line. As was pointed out earlier, in the logistic model, all the six alternative information functions are positive for the entire range of  $\theta$  , while the same is not true for the five alternative information functions in Model B. This result was expected from the result for the basic functions, which were observed earlier in this chapter. It has also been pointed out earlier in this chapter that the item information function always assumes non-negative values for the entire range of  $\theta$  , regardless of the behavior of the alternative information functions, and this is exemplified in these two figures.

The usefulness of the item information function has been emphasized earlier (Samejima, 1977a), especially in connection with the maximum likelihood estimation of the examinee's ability. It should be noted, however, that the blind use of the item information function, or the test information function, is harmful, when the item response information functions, or the alternative information functions, are not always non-negative. This is exemplified in the criticism related with the three-parameter logistic model (Samejima, 1973b). With models of higher complexities, like Model B, care should be taken in finding out the limitation in using the item information function.

As the logical consequence of the observations made earlier for the basic functions, we find that for a certain interval of  $\theta$ , which covers higher levels, the item information function in Model B is practically identical with the corresponding item information function in the logistic model. We can see in Figure 4-6 that this interval is approximately  $(0.4, \infty)$  for item 1, and in Figure 4-7 that it is approximately  $(0.8, \infty)$  for item 2. It is also noted that for these intervals each alternative information function is practically identical with the counterpart in the logistic model, the fact which indicates that the effect of noises caused by random guessing is negligibly small in these intervals, and, therefore, we can expect that the accuracy of ability estimation is as high as in the logistic model in these intervals of  $\theta$ .

Because of the impossibility of presenting the two corresponding graphs of the logistic model and of Model B in one figure for the other three test items, the alternative information functions and the item information function in the logistic model for items 3, 4 and 5 are presented in Figures 4-8 through 4-11, and those in Model B are shown in Figures 4-12 through 4-14.

In each figure for Model B, the value of the parameter  $b_1$  is pointed out by an arrow. We can see that the difference between the two sets of alternative information functions is enhanced, as the discrimination parameter,  $a_g$ , becomes greater.

As an additional information, the basic functions of four hypothetical, binary items following the logistic model and those in the three-parameter logistic model are presented in Appendix II,

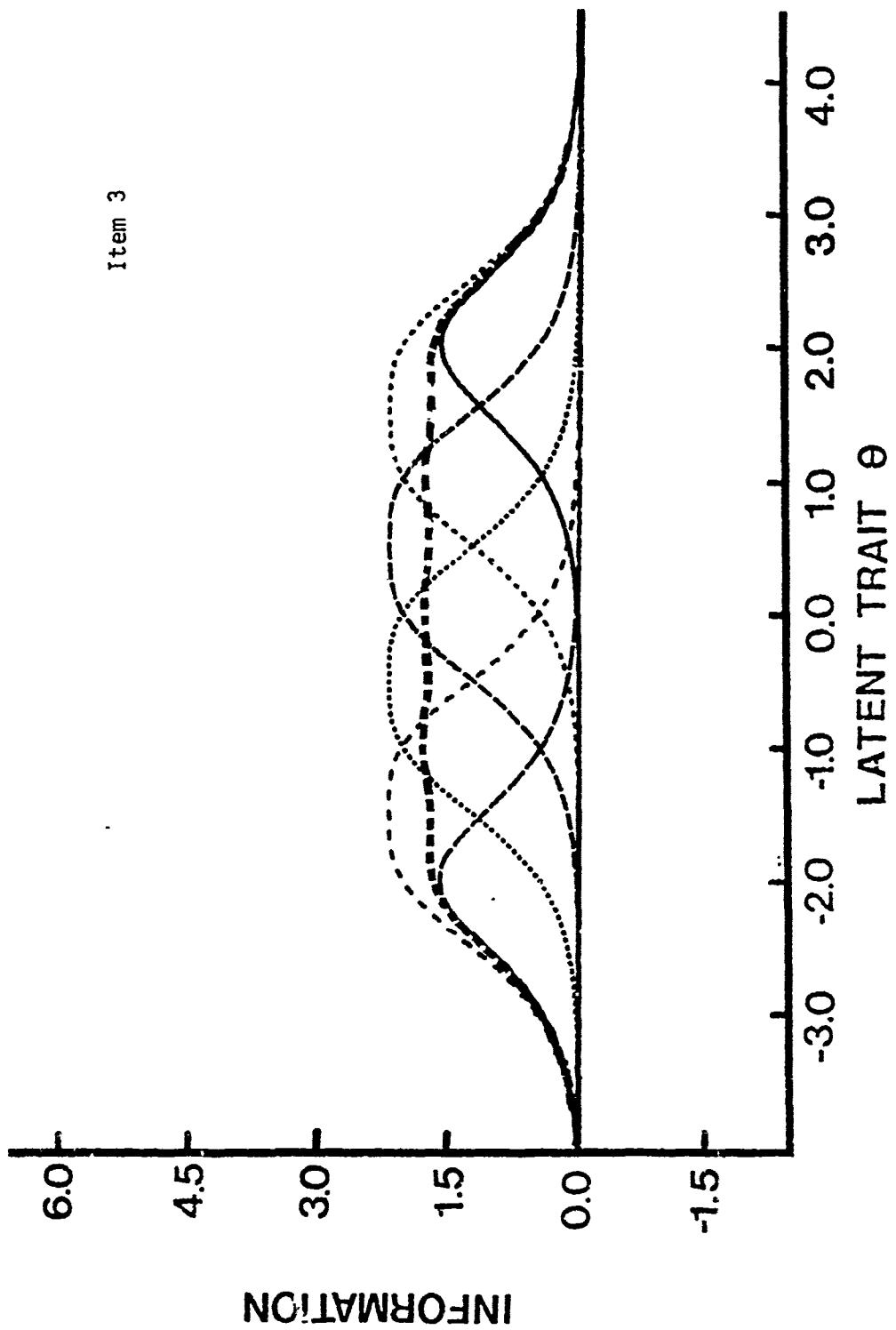


FIGURE 4-8

Item Response Information Functions of Six Response Categories (Various Thinner Curves) and the Item Information Function (Heavy Dashes) in the Logistic Model. The Item Parameters Are:  $a_g = 1.50$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .

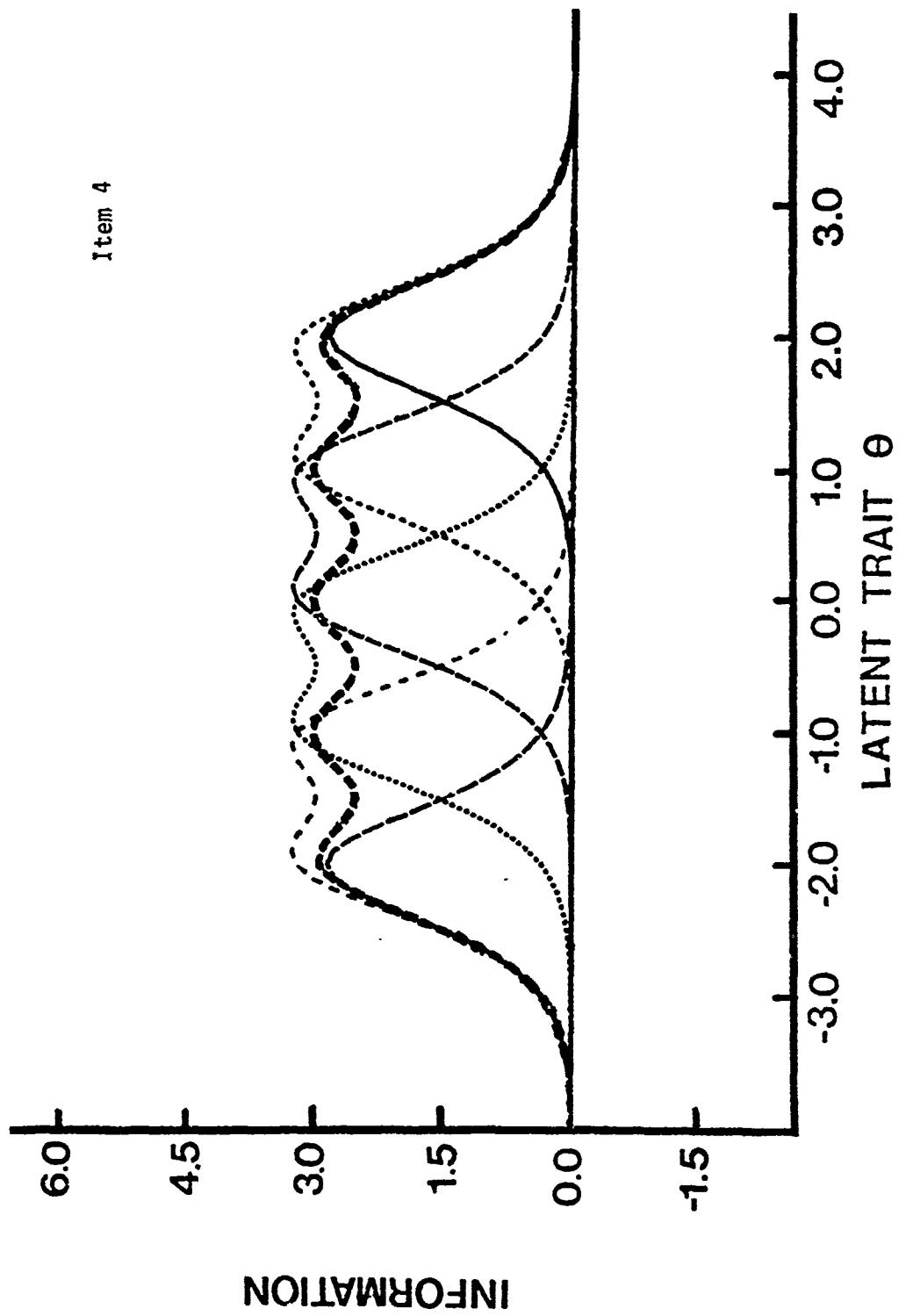
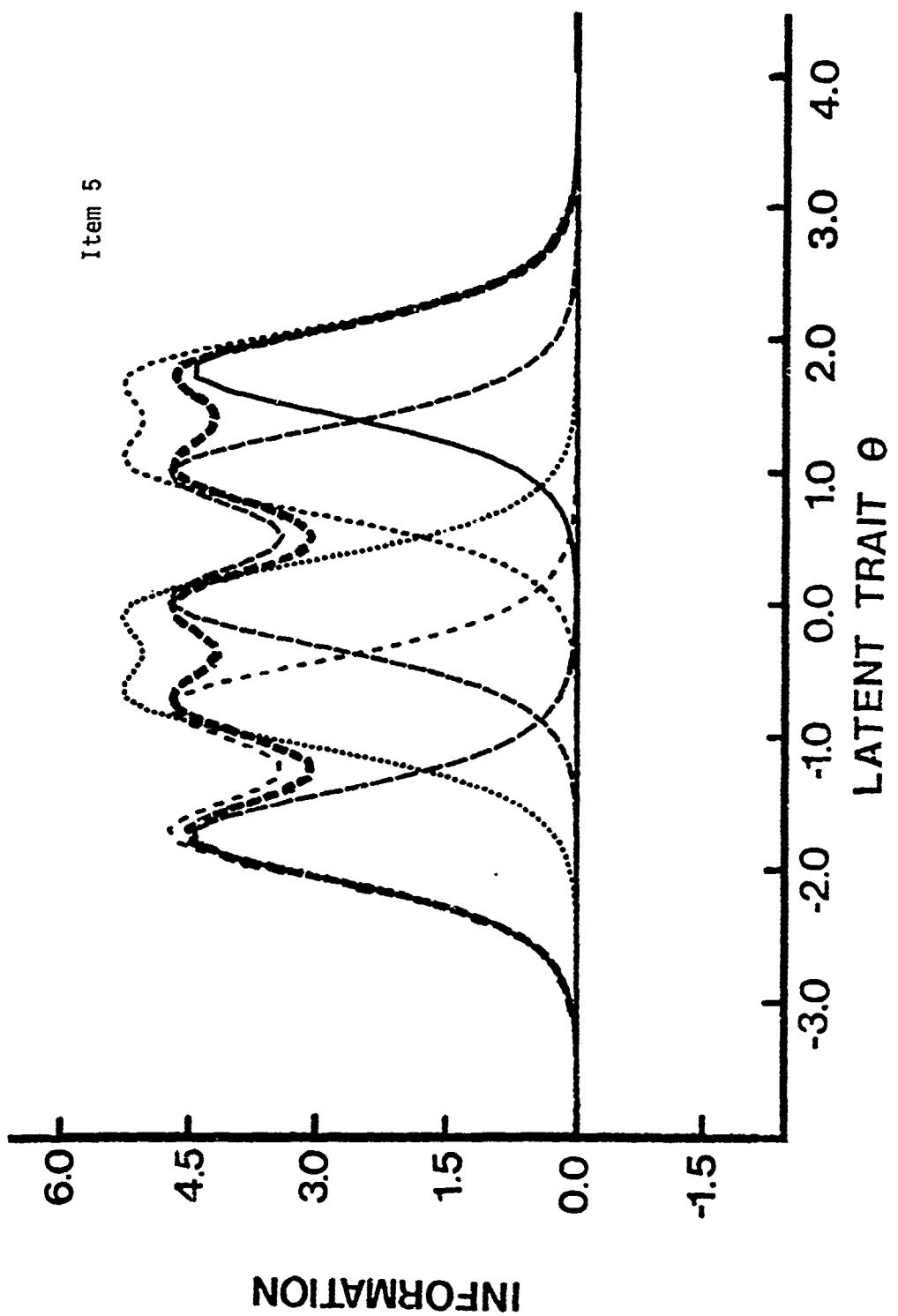


FIGURE 4-9

Item Response Information Functions of Six Response Categories (Various Thinner Curves) and the Item Information Function (Heavy Dashes) in the Logistic Model. The Item Parameters Are:  $a_8 = 2.00$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .



Item Response Information Functions of Six Response Categories (Various Thinner Curves) and the Item Information Function (Heavy Dashes) in the Logistic Model. The Item Parameters Are:  $a_g = 2.50$ ,  $b_1 = -1.75$ ,  $b_2 = -0.75$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 1.75$ .

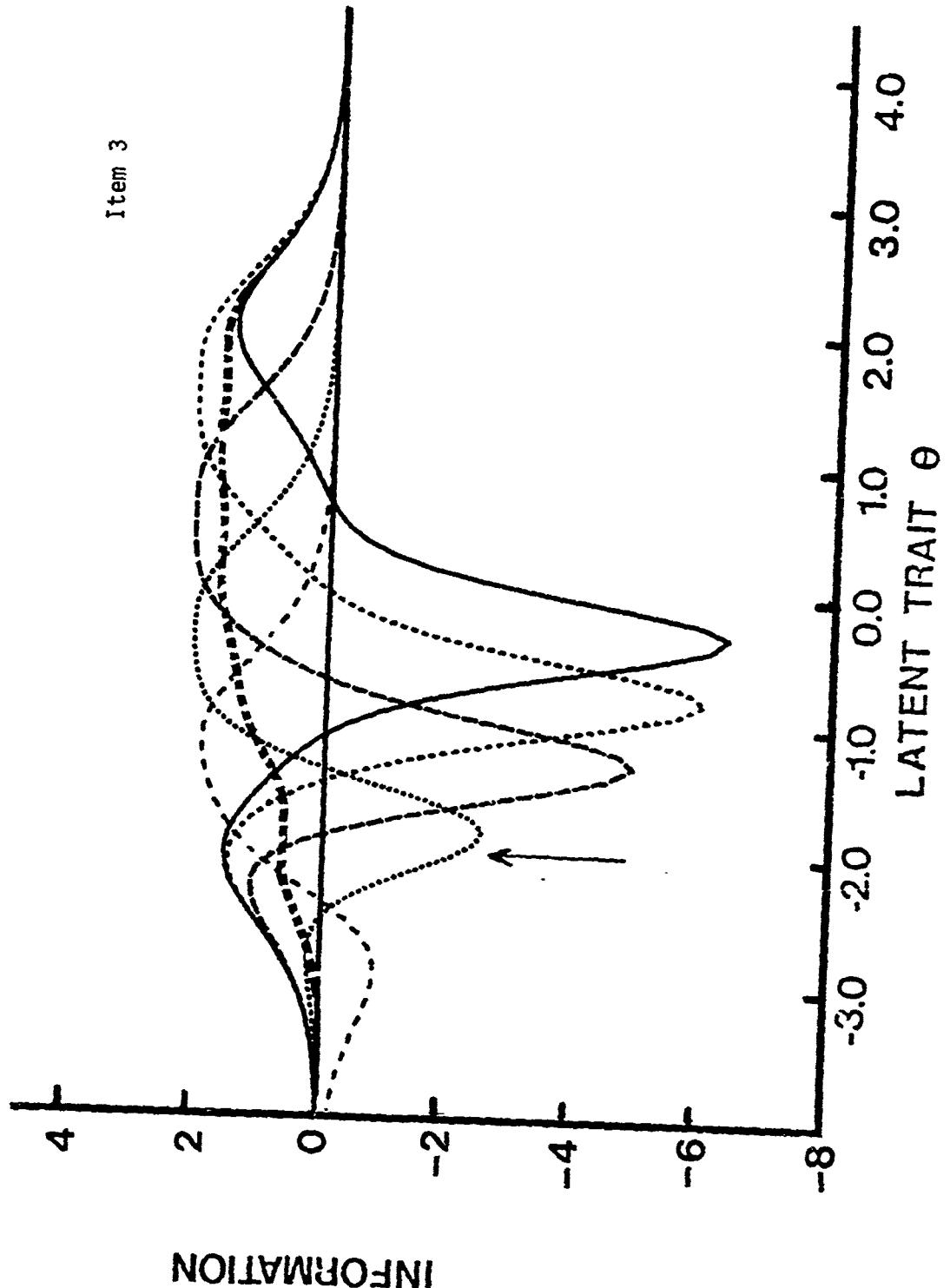


FIGURE 4-11

Item Response Information Functions of Five Alternatives (Various Thinner Curves) and the Item Information Function (Heavy Dashes) in Model B of Type I for the Multiple-Choice Item. The Item Parameters Are:  $a_g = 1.50$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .

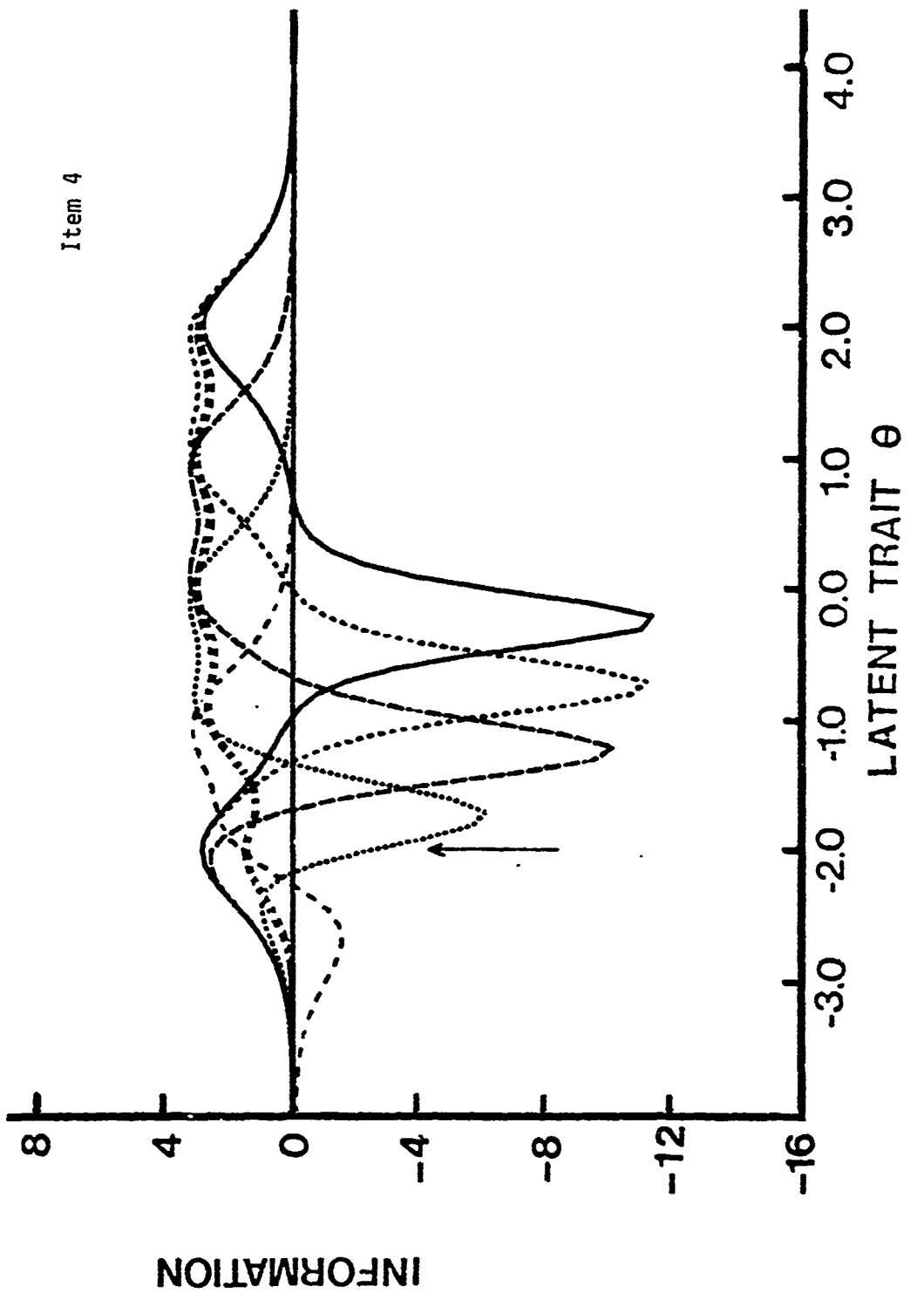


FIGURE 4-12  
Item Response Information Functions of Five Alternatives (Various Thinner Curves) and the Item Information Function (Heavy Dashes) in Model B of Type I for the Multiple-Choice Item. The Item Parameters Are:  $a_g = 2.00$ ,  $b_1 = -2.00$ ,  $b_2 = -1.00$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 2.00$ .

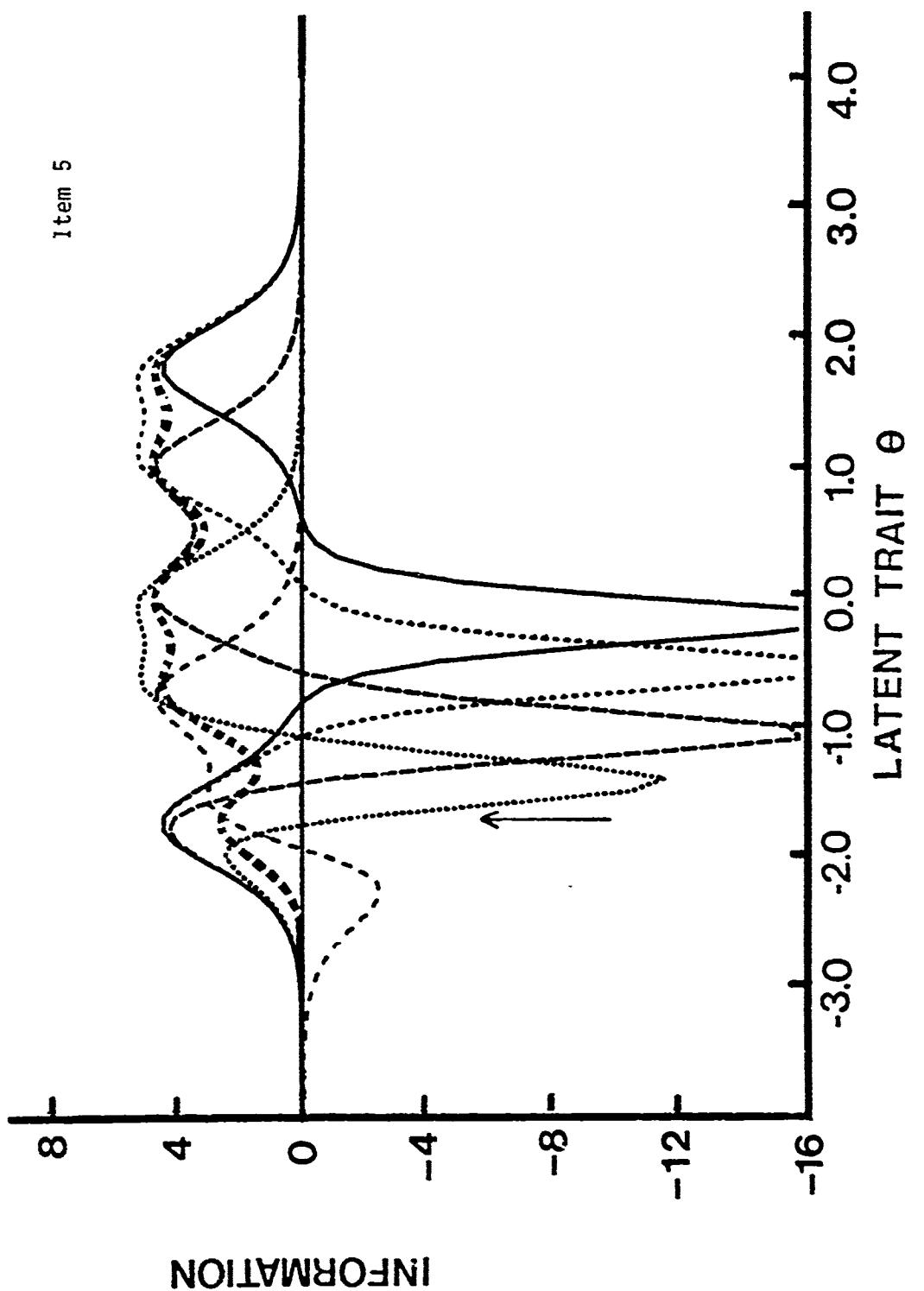


FIGURE 4-13

Item Response Information Functions of Five Alternatives (Various Thinner Curves) and the Item Information Function (Heavy Dashes) in Model B of Type I for the Multiple-Choice Item. The Item Parameters Are:  $a_g = 2.50$ ,  $b_1 = -1.75$ ,  $b_2 = -0.75$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 1.75$ .

as Figures A-2-1 and A-2-2, with the discrimination parameters  $a_g = 1.00, 1.50, 2.00$  and  $2.50$ , respectively, and the common set of the difficulty parameter and the guessing parameter,  $b_g = 0.00$  and  $c_g = 0.20$ . If we compare them with those results of Model B, it is obvious that we should expect a substantially different outcome resulting from the analysis of data following Model B, from the one obtained by adopting the three-parameter logistic model.

## V Qualities That Distinguish Good Test Items from Bad Ones

Here we stop and think of the qualities that make test items good ones, and how they are related with mental test theory. First of all, there is no question that good test items are those which possess both valid contents and predictability, whose relative weights depend upon the testing purposes and the complexities of performances involved. Secondly, good test items are informative items, in the sense that they provide us with an accurate discrimination of individuals with respect to the ability measured. This second major point is closely related with mental test theory, and it is the theorist's contribution that makes it possible to extract valuable information from the test result, and to estimate the individual's ability efficiently. In so doing, however, essential considerations resulting from theory must be understood by the test constructors, and taken into account in the early stage of test development.

There is no doubt that one of the strongest negative factors involved in the multiple-choice item is the noise caused by random guessing. If the examinee's behavior follows the knowledge or random guessing principle and the three-parameter normal ogive, or logistic, model, for example, the multiple-choice item is nothing but a "blurred" image of the free-response item, with larger errors of ability estimation especially on lower levels of ability. This is still the case when we make better use of the multiple-choice item so that it is no longer a poor substitute for the free-response item, but is a type of test item which has a high potential for

being highly informative. We can say, therefore, that good multiple-choice items are those which have little room for noises to contaminate the meaning of correct answers.

We have seen in Chapter 3 that, if the examinee's behavior follows, say, Model A of Type I for the multiple-choice item, then under certain conditions the operating characteristic of the correct answer is practically the same as that of the normal ogive model on the graded response level, which includes no random guessing effect, except for the "tail" which lies on lower levels of ability. We notice that this situation is materialized when the discrimination parameter,  $a_g$ , is high, and the first difficulty parameter,  $b_1$ , is substantially lower compared with the  $m$ -th difficulty parameter,  $b_m$ . This second condition implies that we need a distractor which attracts examinees whose ability levels are lowest among other examinees of the group in which we are interested. In such a case, ability estimation is as efficient as the one provided by a free-response item, and with additional information given by the distractors it can be substantially better, as far as our examinee group is concerned. Effort should be put, therefore, upon finding an answer which does not attract people of high ability, but is appealing to those whose ability is lower.

An interesting, and valuable, by-product of the above effort is that we can use such a set of multiple-choice items for investigating the operating characteristics of their own distractors. We recall that the methods and approaches for estimating the operating characteristics without assuming any mathematical forms (cf. Samejima,

1977c, 1978a through 1978f) have two common restrictions, such that: 1) we need a set of test items, or Old Test, whose operating characteristics are known; and 2) the test information function of the Old Test must be constant and substantially high. It can be shown that the second restriction is not really a restriction, and we can make use of a test whose information function is not constant. (This will be shown and discussed in a separate paper.) The necessity of the Old Test is more restrictive under the general circumstances. Suppose we have a set of a reasonably large number of multiple-choice items, which are well constructed and each of which contains a good, i.e. level distracter, so that the operating characteristics of the correct answers are practically the same as those without the effect of noises caused by random guessing, for the levels of ability where the examinees in question are located. We score each of these items as a binary item, and make the binary response pattern for each examinee. These response patterns can be used as the substitutes for the response patterns based on the Old Test, and the maximum likelihood estimate of each examinee's ability can be obtained. Then we divide the total group of examinees into m subgroups, in accordance with their alternative selection of the specific item. The operating characteristic of each distracter can be estimated by using the set of maximum likelihood estimates obtained from the binary response patterns in one of the combinations of method and approach for estimating the operating characteristics. We need only one set of test items, therefore, instead of two, in estimating the operating characteristics of the

distractors.

Another important consideration in the effort of constructing good test items is that we select alternatives for each item in such a way that the resulting random guessing by examinees occurs only when they have no idea about the correct answer. We recall that this is one of the essential parts of our rationale behind the family of models for the multiple-choice item, which is proposed in this paper. We notice that each item has its own set of possible answers, or answer space. For the purpose of illustration, suppose that our question is  $3 + 5 = ?$ . In this example, it is highly unlikely that the examinee conceives of "house" as the answer. The same is true with any other words, or numbers such as  $3.14159$ ,  $-27$ ,  $\sqrt{28915}$ ,  $604/917$ , etc., or even integers like  $300$ ,  $1,023$ , etc. Thus we can assume there is some common answer space for most of the individuals who have a possibility of taking the test in question. For this specific item, this common answer space may be the set of all the positive integers which do not exceed  $19$ . Suppose that our selection of the alternatives for this test item is  $7$ ,  $10$ ,  $6$  and  $12$ , in addition to the correct answer,  $8$ . If an examinee is convinced that the correct answer is  $9$ , then he will be confused because he cannot find his answer among the alternatives, and, because of his confusion, he may end up with selecting one of the five alternatives randomly. In this example, the examinee does have some clear idea about the correct answer to the question, and he should not be categorized into the lowest item score group, i.e., the "no idea at all." If we do, the principle behind the family of models will

be violated, and the application of one of the models in the family will produce nothing but artifacts, which are a "distorted image" of the psychological reality.

The above example illustrates the importance of the selection of alternatives for the multiple-choice item. We must put our effort, therefore, upon selecting a set of alternatives in such a way that they can be distractors which attract separate groups of examinees whose ability levels are centered into successively arranged subintervals of ability. This is rather difficult to realize for numerical items like the one exemplified in a previous paragraph, but is less difficult for, say, verbal items. Considering that the scoring of free-response, numerical items is, in general, much easier and the multiple-choice item is not needed as much in this area, this fact is not a serious obstacle.

Thus we have seen that one of the criteria for good multiple-choice items is that they possess distractors which most strongly appeal to the separate groups of examinees whose abilities are located at different subintervals of the ability continuum and, together with the one for the correct answer, they cover a certain interval of ability exhaustively. Furthermore, the lowest of such subintervals of ability which is essentially attracted by one distractor must be substantially lower in comparison with the corresponding subinterval for the correct answer, so that the operating characteristic of the correct answer is "untouched" by the effect of noises caused by random guessing.

One logical consequence of developing such good multiple-choice items is that we can use the items for tailored testing, and the selection of an alternative for a specific item provides us with a good way of branching examinees. It can easily be seen that good multiple-choice items, which were discussed in the preceding paragraph, are efficient items for such a branching in tailored testing.

The discussion developed in the present chapter will become more meaningful and productive, if it is extended to include real data. This part of research will be conducted in the near future.

## VI Discussion and Conclusions

A new family of models for the multiple-choice item, which includes Models A, B and C of Type I, was proposed. The basic function of the alternative, and the alternative information function, are discussed mainly in Model B, which is closely related with the logistic model, in addition to the operating characteristic of the alternative. Qualities which distinguish good multiple-choice items from bad ones were discussed, in the light of the new family of models.

The author believes that this family of models explains the psychological reality much better than any other model built for the multiple-choice item. Further investigation of the characteristics of each model of Type I is essential, however, since the present paper is just the beginning and various considerations should be made before we can make full use of the model and of multiple-choice items. Among others, we need to pursue the limitation and restriction which should be put upon the use of the item information function as the measure of accuracy in ability estimation, because of the complexities of the models in this family.

The traditional use of the three-parameter logistic model for the multiple-choice item must be terminated, unless someone can present a strong theoretical support. Empirical support, such as the one given by Lord (Lord, 1970) cannot be strong enough, since we can conceive of many other mathematical forms that will fit the data just as well. It has been reported by

many researchers (e.g., Lord, 1968) that the estimated guessing parameter,  $c_g$ , is less than the reciprocal of the number of alternatives. The parameter is meaningful only when the rationale behind the three-parameter logistic model is acceptable, and the above fact itself disproves the rationale as far as these data are concerned.

In relation with the above fact, it should be emphasized that researchers engage themselves in model validation whenever they use any models. The blind use of mathematical models will create nothing but disaster, and will hinder the progress of science.

An emphasis should also be put upon the scientific attitude of perceiving the object as it is. Although the characteristic of the multiple-choice item as a substitute for the free-response item is only a part of its nature, most researchers have been blind enough to ignore the information given by distractors, which the free-response item is not able to provide.

It is interesting to note that, in the history of multiple-choice tests, test constructors have more or less depended upon their intuition and included distractors in the set of alternatives in their effort of developing good test items, while data analysts have completely ignored the information which can be obtained from the distractors, by depending upon the three-parameter logistic, or normal ogive, model. This contradiction caused by the two different attitudes of two different groups of people has had an unfortunate, negative effect on the progress of this area of science,

and, evidently, the blame should be put upon the data analysts.

We must say, therefore, that researchers are indebted to the test constructors who have produced good multiple-choice items, and they should make up for the "lost time" by trying hard to change their orientation. For this purpose, the family of models presented in this paper may serve well, taking the place that three-parameter models have occupied for so long. It may change the direction of tailored testing drastically, for we can use the information given by distractors for branching the examinees, in the way that already realized by Shiba without depending upon the present family of models (cf. Samejima, 1980).

REFERENCES

- [1] Bejar, I. I., D. J. Weiss and G. G. Kingsbury. Calibration of an item pool for the adaptive measurement of achievement. Office of Naval Research Report 77-5, 1977.
- [2] Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord and M. R. Novick; Statistical theories of mental test scores. Addison-Wesley, 1968, Chapters 17-20.
- [3] Bock, R. D. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. Psychometrika, 37, 1972, pages 29-51.
- [4] Hambleton, R. K. and J. Gifford. Robustness of latent trait models. Paper presented at the 1979 Computerized Adaptive Testing Conference, Minneapolis, Minnesota, 1979.
- [5] Lord, F. M. An analysis of the verbal scholastic aptitude test using Birnbaum's three-parameter logistic model. Educational and Psychological Measurement, 28, 989-1020, 1968.
- [6] Lord, F. M. Item characteristic curves as estimated without knowledge of their mathematical form -- a confrontation of Birnbaum's logistical model. Psychometrika, 35, 43-50, 1970.
- [7] Lord, F. M. and M. R. Novick. Statistical theories of mental test scores. Addison-Wesley, 1968, Chapter 16.
- [8] McBride, J. R. Some properties of a Bayesian adaptive ability testing strategy. Applied Psychological Measurement, 1, 121-140, 1977.
- [9] Reckase, M. D. Unifactor latent trait models applied to multifactor tests: results and implications. D. J. Weiss (Ed.), Proceedings of the 1977 Computerized Adaptive Testing Conference, 354-372, 1977.
- [10] Samejima, F. Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph, No. 17, 1969.
- [11] Samejima, F. A general model for free-response data. Psychometrika Monograph, No. 18, 1972.
- [12] Samejima, F. Homogeneous case of the continuous response model. Psychometrika, 1973a, 38, 203-219.
- [13] Samejima, F. A comment on Birnbaum's three-parameter logistic model in the latent trait theory. Psychometrika, 38, 221-233, 1973b.
- [14] Samejima, F. A use of the information function in tailored testing. Applied Psychological Measurement, 1, 1977a, pages 233-247.

REFERENCES (continued)

- [15] Samejima, F. A method of estimating item characteristic functions using the maximum likelihood estimate of ability. Psychometrika, 42, 1977b, pages 163-191.
- [16] Samejima, F. Estimation of the operating characteristics of item response categories I: Introduction to the Two-Parameter Beta Method. Office of Naval Research, Research Report 77-1, 1977c.
- [17] Samejima, F. Estimation of the operating characteristics of item response categories II: Further development of the Two-Parameter Beta Method. Office of Naval Research, Research Report 78-1, 1978a.
- [18] Samejima, F. Estimation of the operating characteristics of item response categories III: The Normal Approach Method and the Pearson System Method. Office of Naval Research, Research Report 78-2, 1978b.
- [19] Samejima, F. Estimation of the operating characteristics of item response categories IV: Comparison of the different methods. Office of Naval Research, Research Report 78-3, 1978c.
- [20] Samejima, F. Estimation of the operating characteristics of item response categories V: Weighted Sum Procedure in the Conditional P.D.F. Approach. Office of Naval Research, Research Report 78-4, 1978d.
- [21] Samejima, F. Estimation of the operating characteristics of item response categories VI: Proportioned Sum Procedure in the Conditional P.D.F. Approach. Office of Naval Research, Research Report 78-5, 1978e.
- [22] Samejima, F. Estimation of the operating characteristics of item response categories VII: Bivariate P.D.F. Approach with Normal Approach Method. Office of Naval Research, Research Report 78-6, 1978f.
- [23] Samejima, F. Constant information model: a new, promising item characteristic function. Office of Naval Research, Research Report 79-1, 1979.
- [24] Samejima, F. Introducing some research related with the multiple-choice test item in Japan: toward the validation of mathematical models. Office of Naval Research, 1980, (in preparation).
- [25] Swaminathan, H. and J. Gifford. Precision of item parameter estimation. Paper presented at the 1979 Computerized Adaptive Testing Conference, Minneapolis, Minnesota, 1979.

REFERENCES (continued)

- [26] Sympson, J. B. A model for testing with multidimensional items.  
D. J. Weiss (Ed.), Proceedings of the 1977 Computerized Adaptive Testing Conference, 82-98, 1977.
- [27] Urry, V. W. Tailored Testing: A successful application of latent trait theory. Journal of Educational Measurement, 1977, 14, 181-196.
- [28] Warm, T. A. A primer of item response theory. Coast Guard Technical Report 940279, 1978.

APPENDIX I

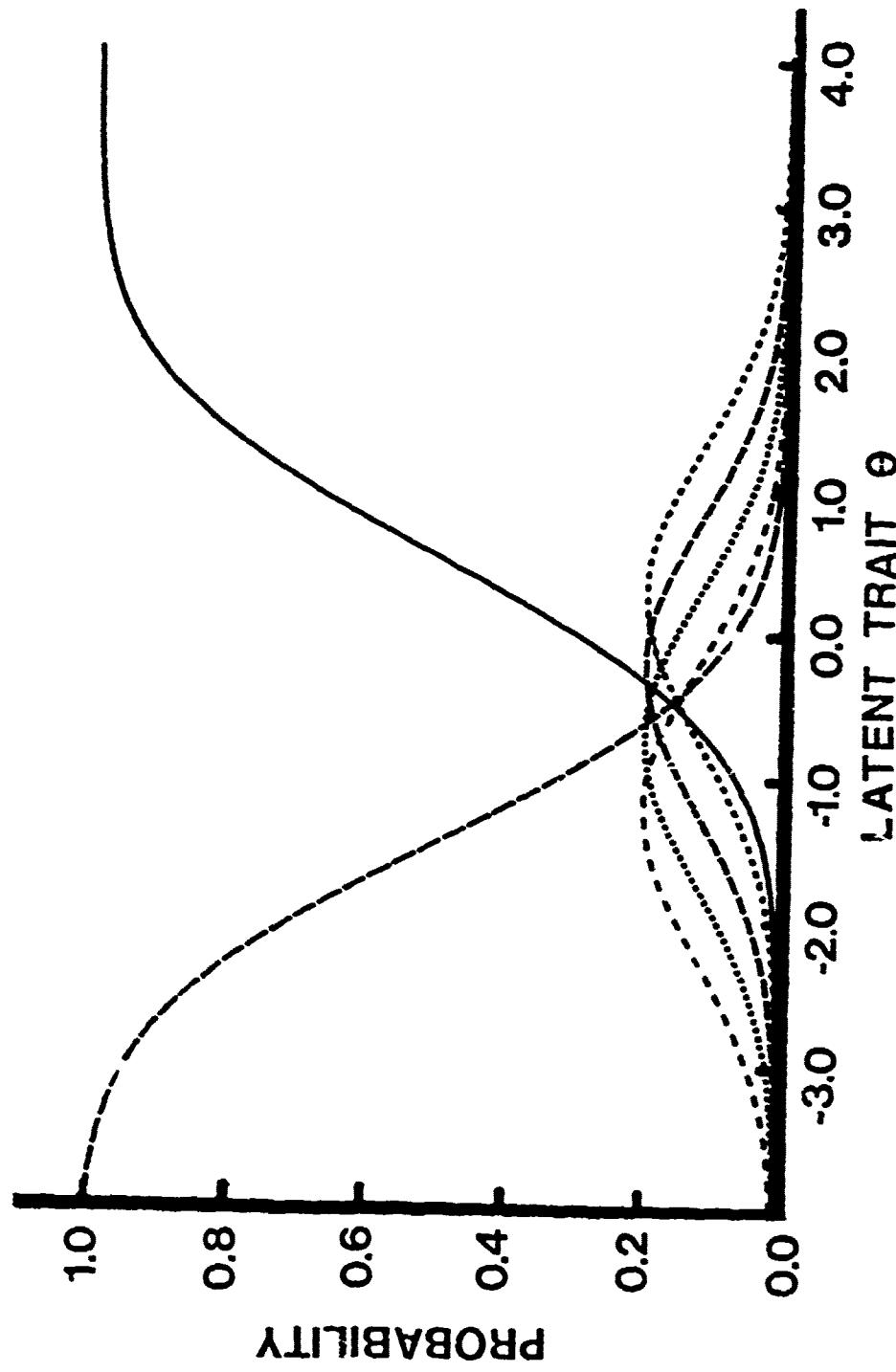


FIGURE A-1-1

Operating Characteristics of Six Item Response Categories Following the Normal Ogive Model, with  $a_8 = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -1.00$ ,  $b_3 = -0.50$ ,  $b_4 = 0.00$  and  $b_5 = 0.50$ .

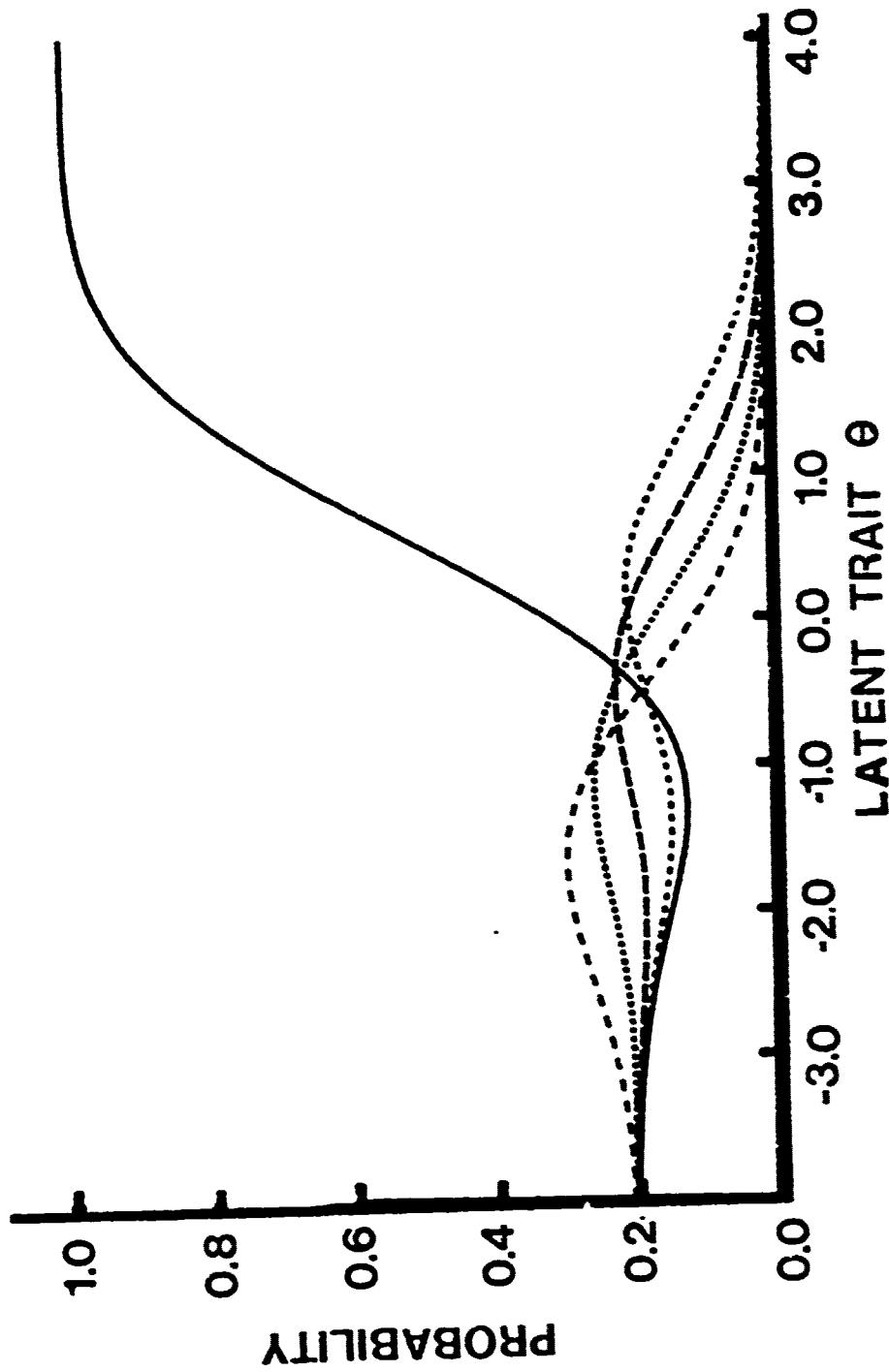


FIGURE A-1-2

Operating Characteristics of Five Alternatives following the Model A of Type I for the Multiple-Choice Item. The Parameters Are:  $a_0 = 1.00$ ,  $b_1 = -1.50$ ,  $b_2 = -1.00$ ,  $b_3 = -0.50$ ,  $b_4 = 0.00$  and  $b_5 = 0.50$ .

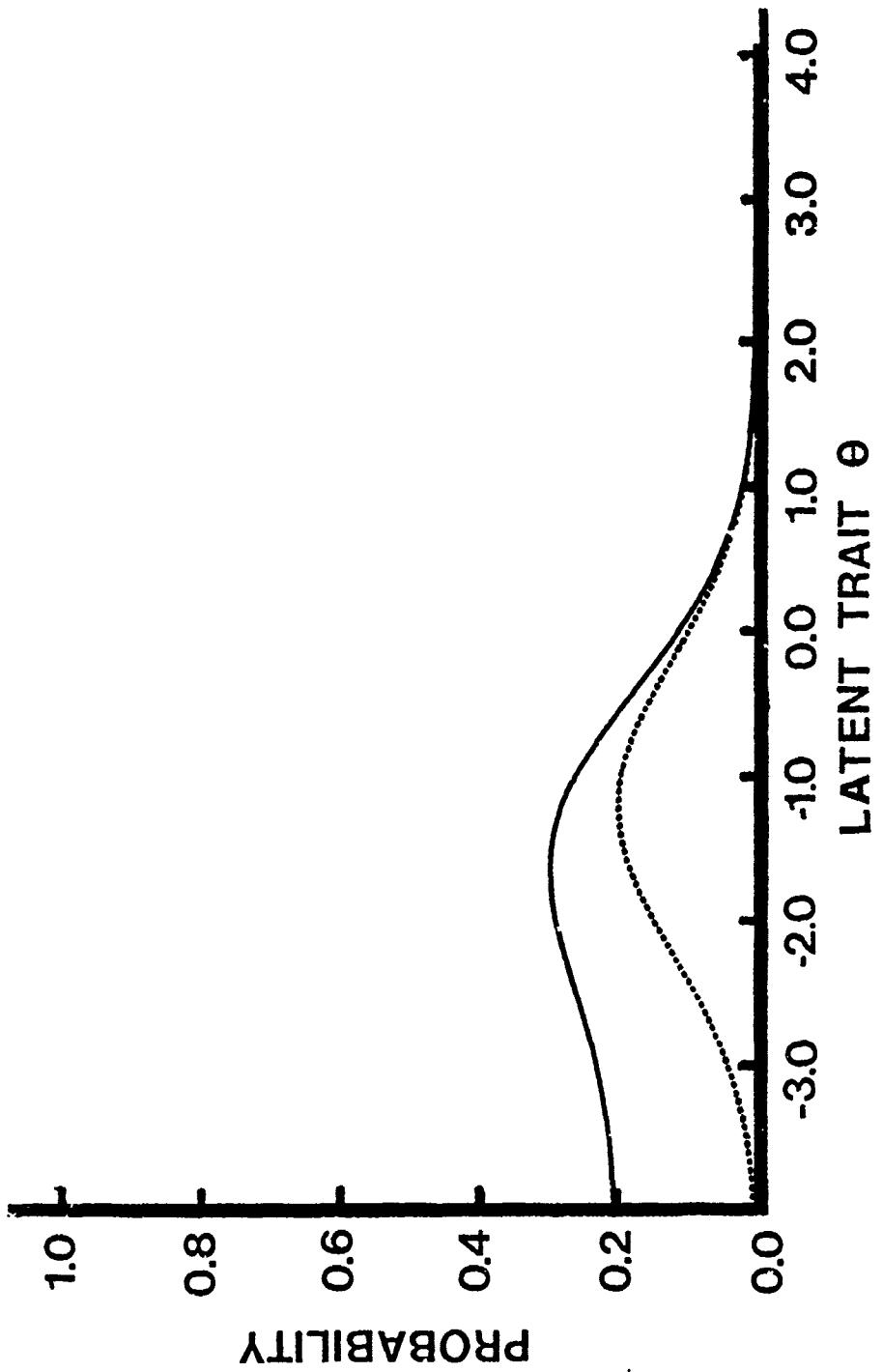


FIGURE A-1-3

Comparison of the Two Operating Characteristics in the Normal Ogive Model (Dotted Curve) and in Model A of Type I (Solid Curve).  
 $h = 1$ .

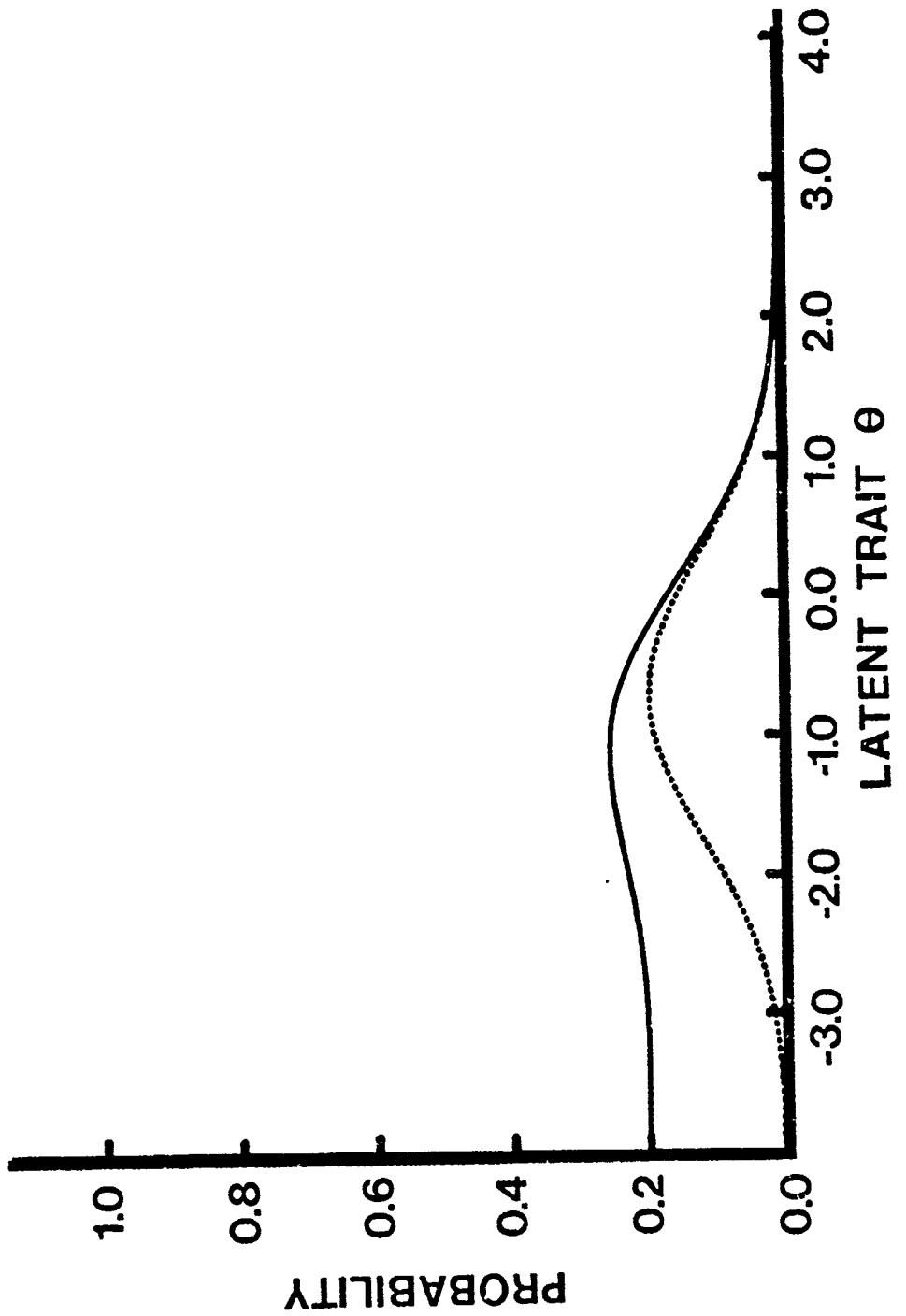


FIGURE A-1-3 (Continued)  $h = 2$ .

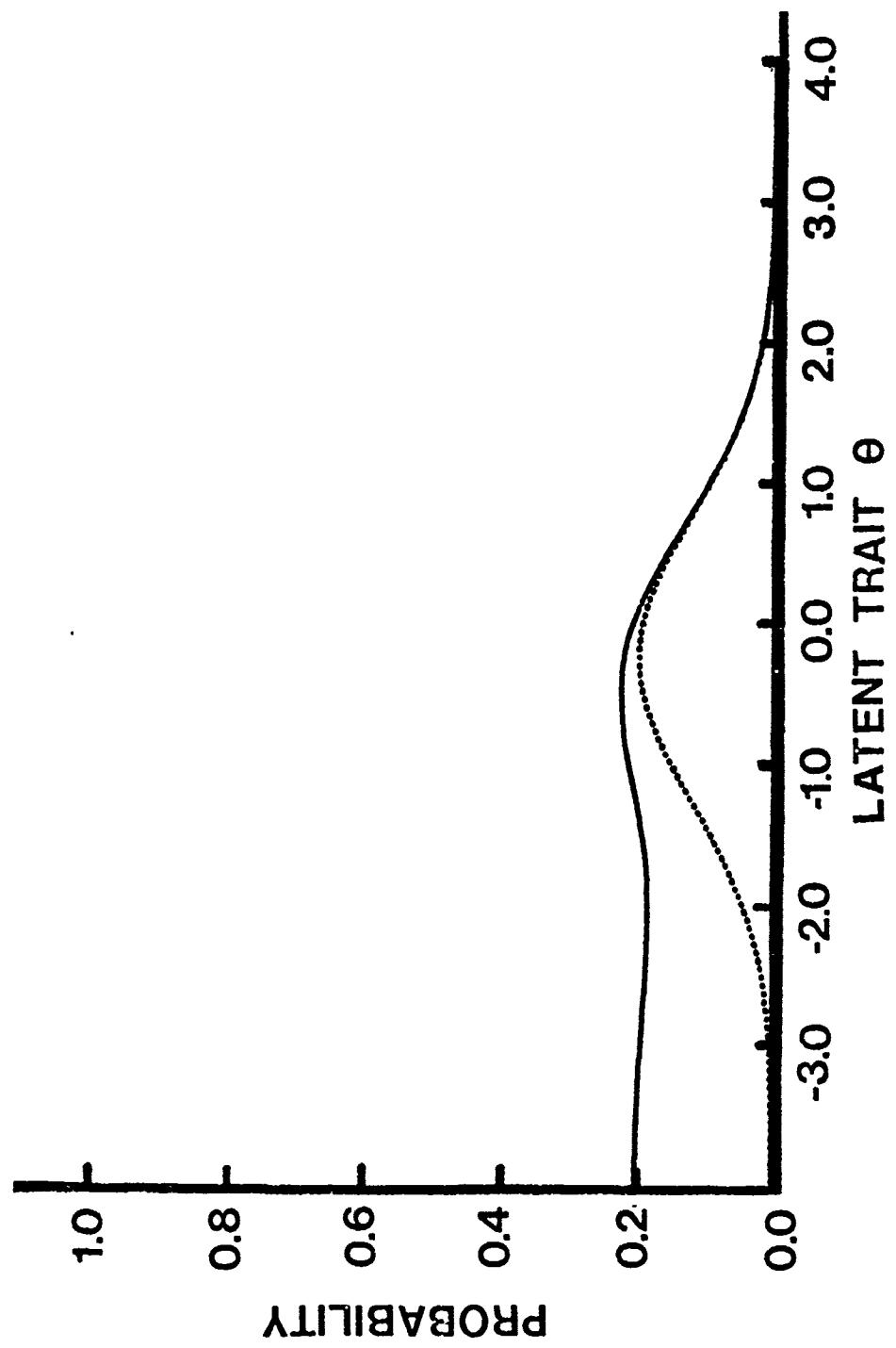


FIGURE A-1-3 (Continued)  $h = 3$ .

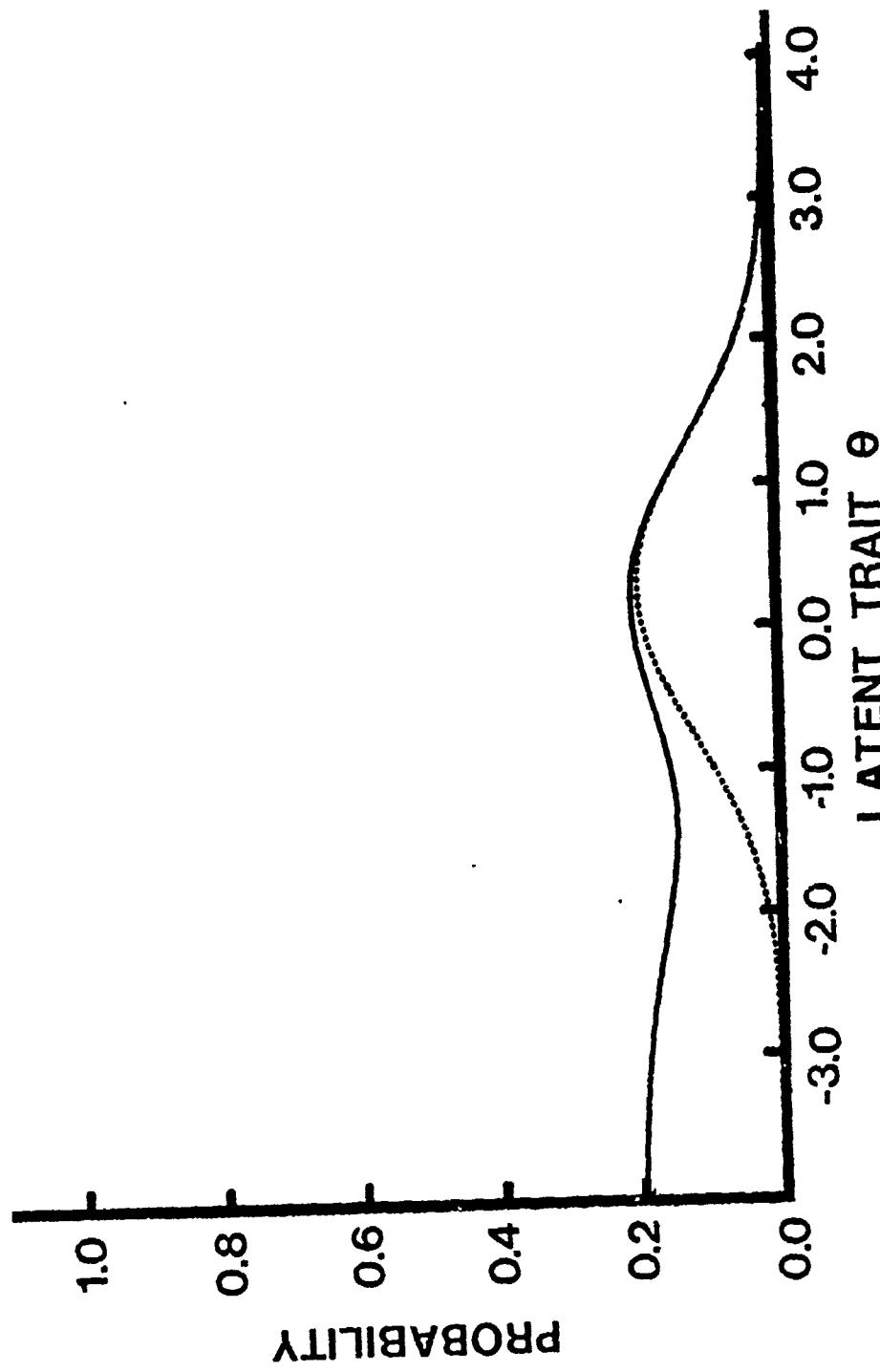


FIGURE A-1-3 (Continued)  $h = 4$ .

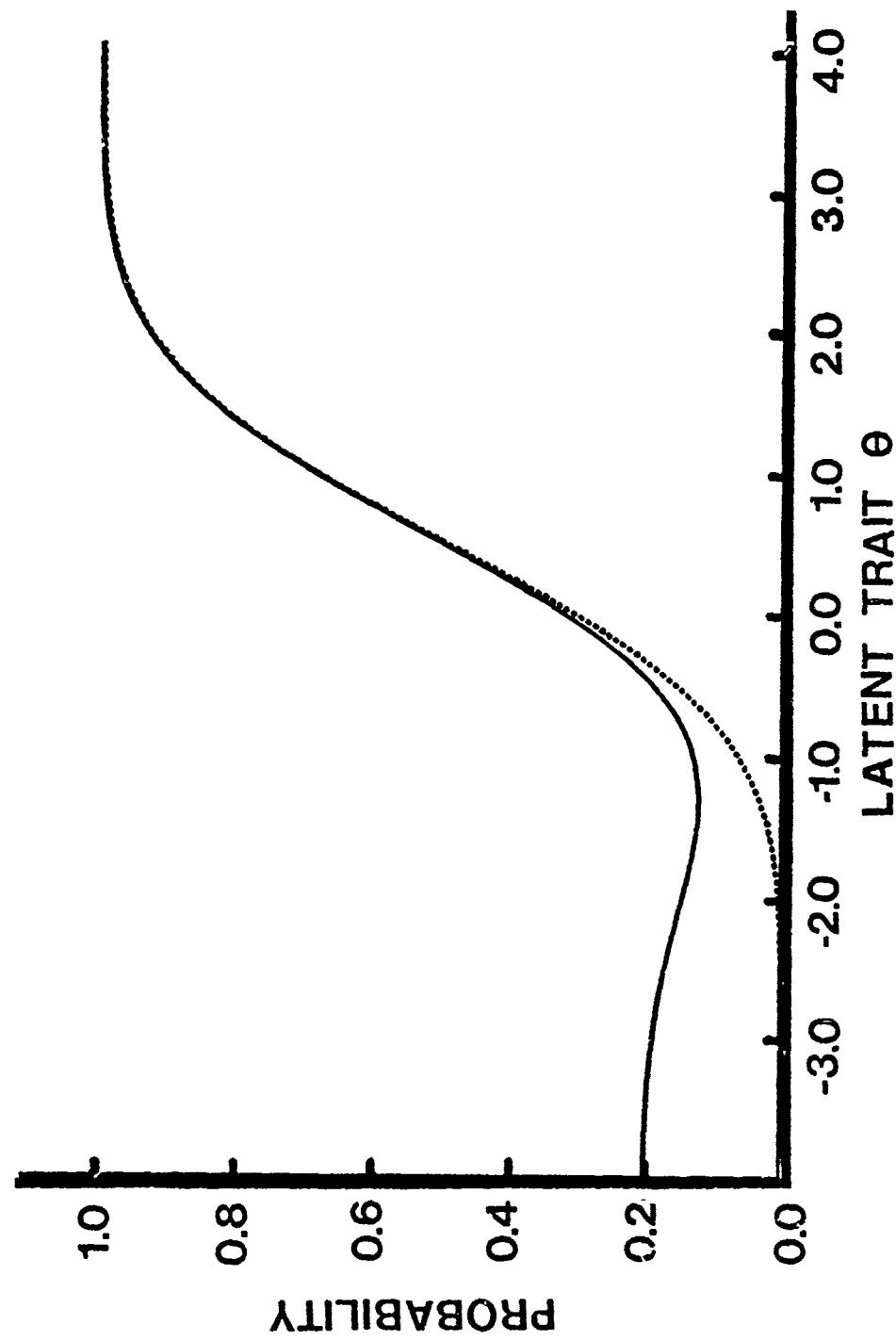


FIGURE A-1-3 (Continued)  $h = 5$ .

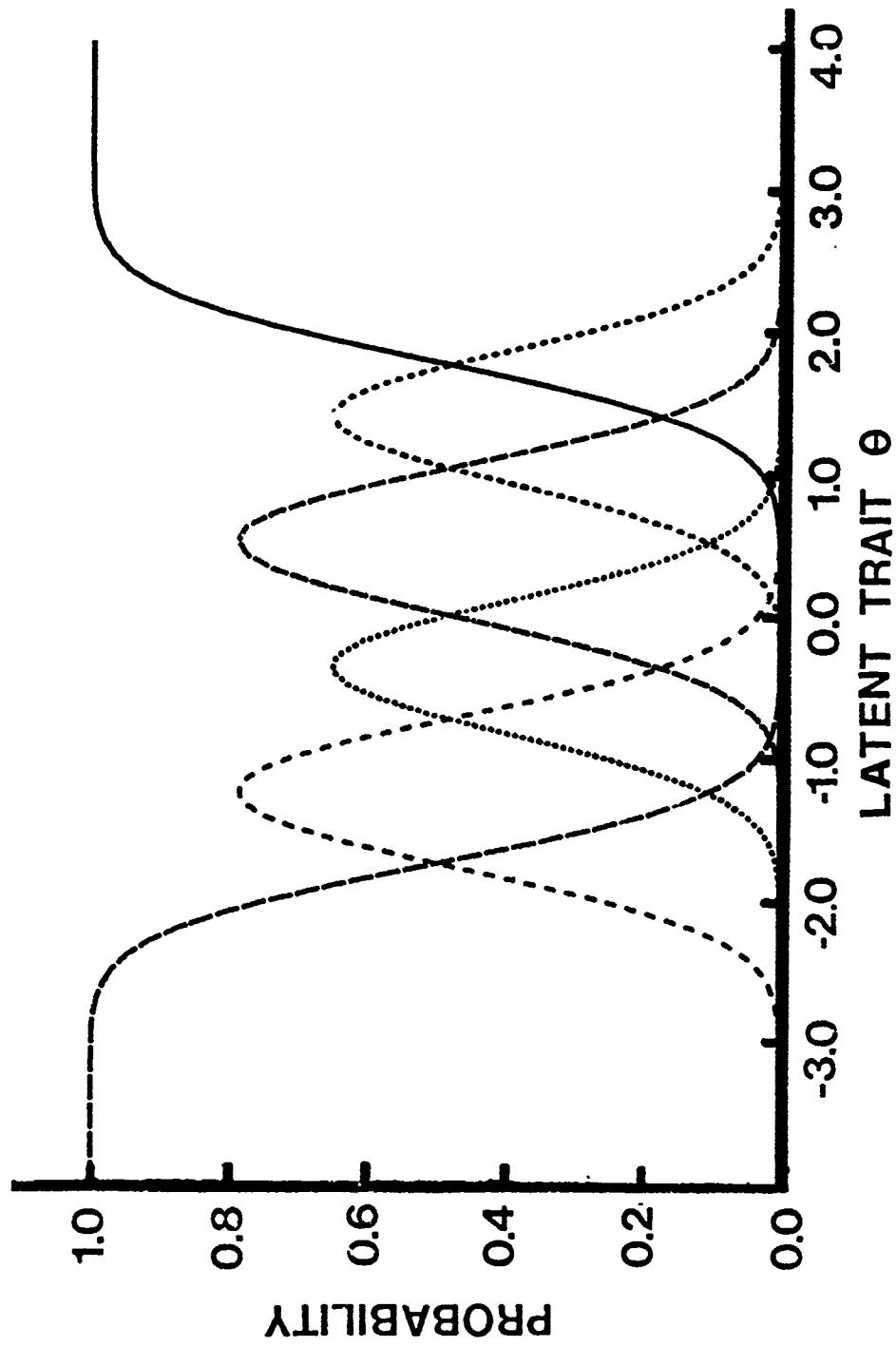


FIGURE A-1-4

Operating Characteristics of Six Item Response Categories Following the Normal Ogive Model, with  $a_g = 2.50$ ,  $b_1 = -1.75$ ,  $b_2 = -0.75$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 1.75$ .

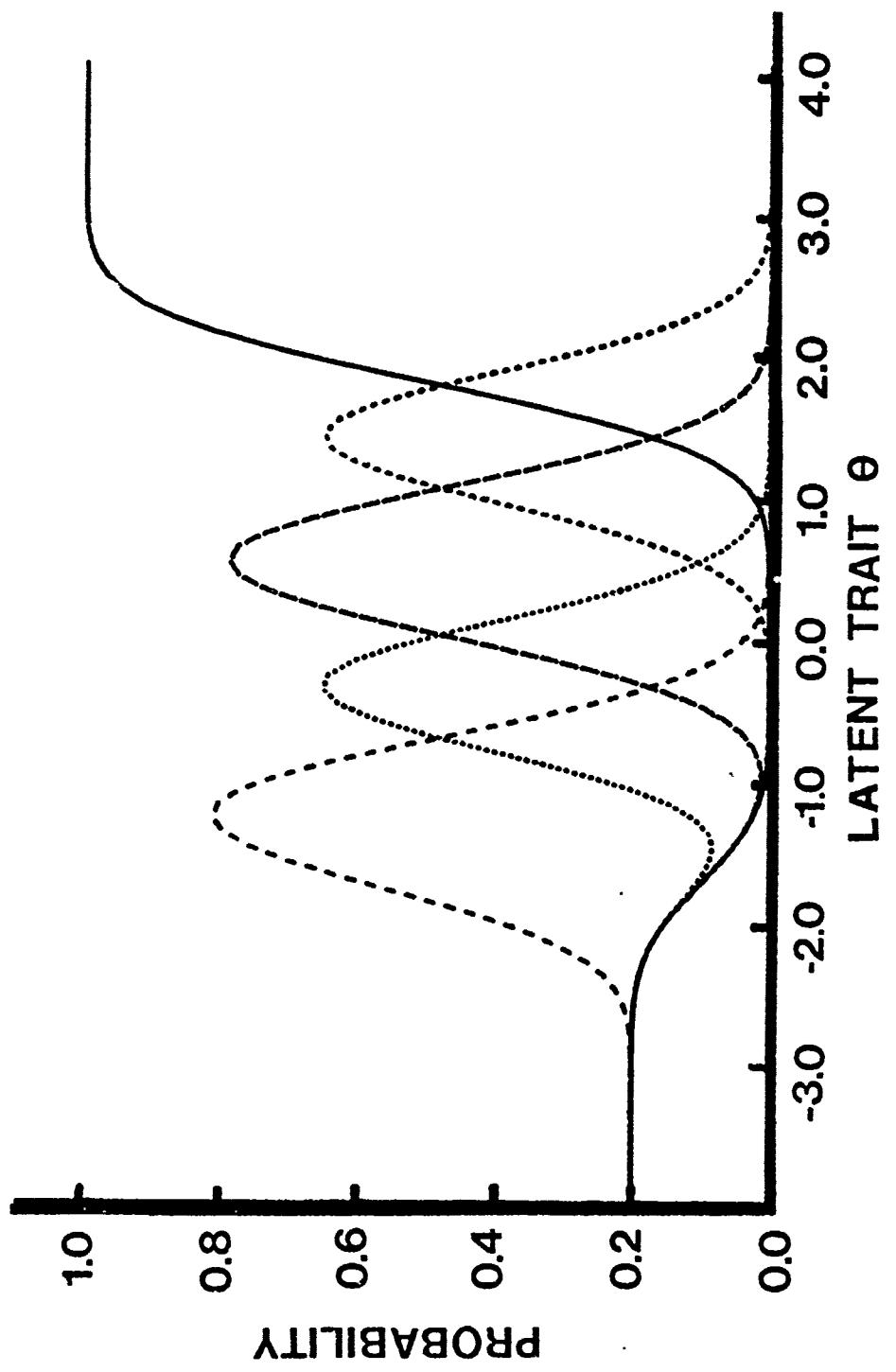


FIGURE A-1-5

Operating Characteristics of Five Alternatives Following the Model A of Type I for the Multiple-Choice Item. The Parameters Are:  $a_g = 2.50$ ,  $b_1 = -1.75$ ,  $b_2 = -0.75$ ,  $b_3 = 0.00$ ,  $b_4 = 1.00$  and  $b_5 = 1.75$ .

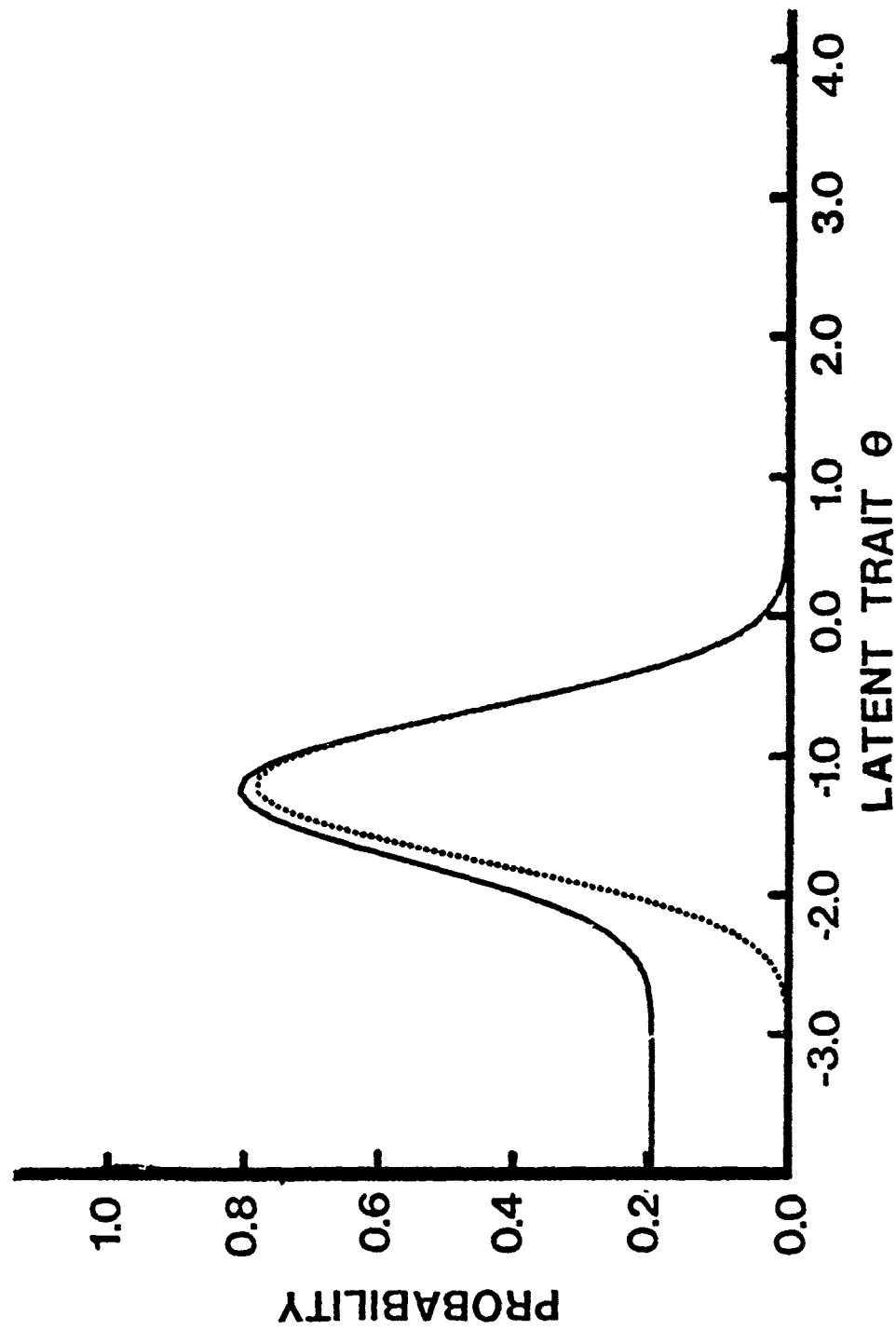


FIGURE A-1-6

Comparison of the Two Operating Characteristics in the Normal Ogive Model (Dotted Curve) and in Model A of Type I (Solid Curve).  $n = 1$ .

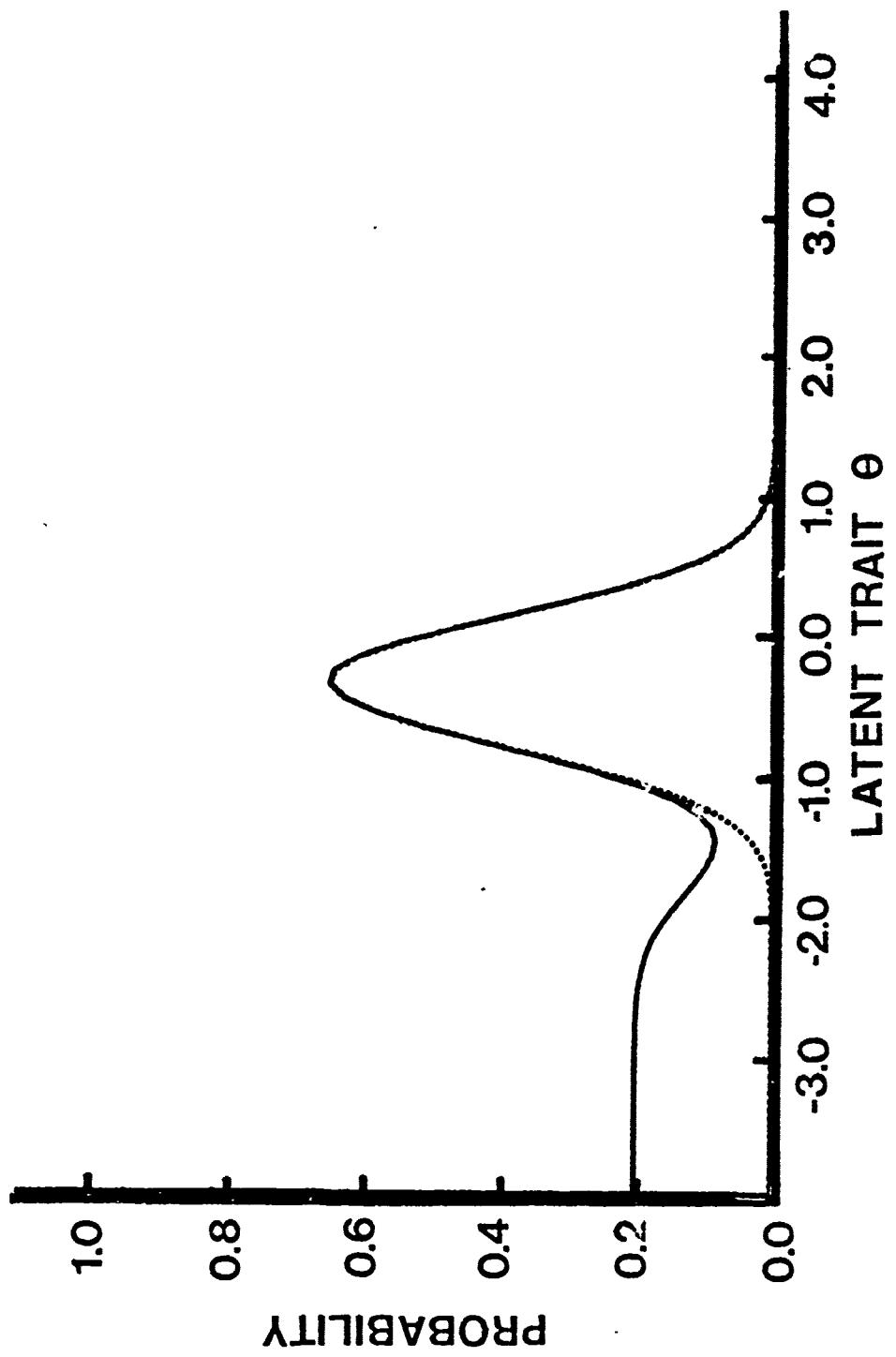


FIGURE A-1-6 (Continued)  $h = 2$ .

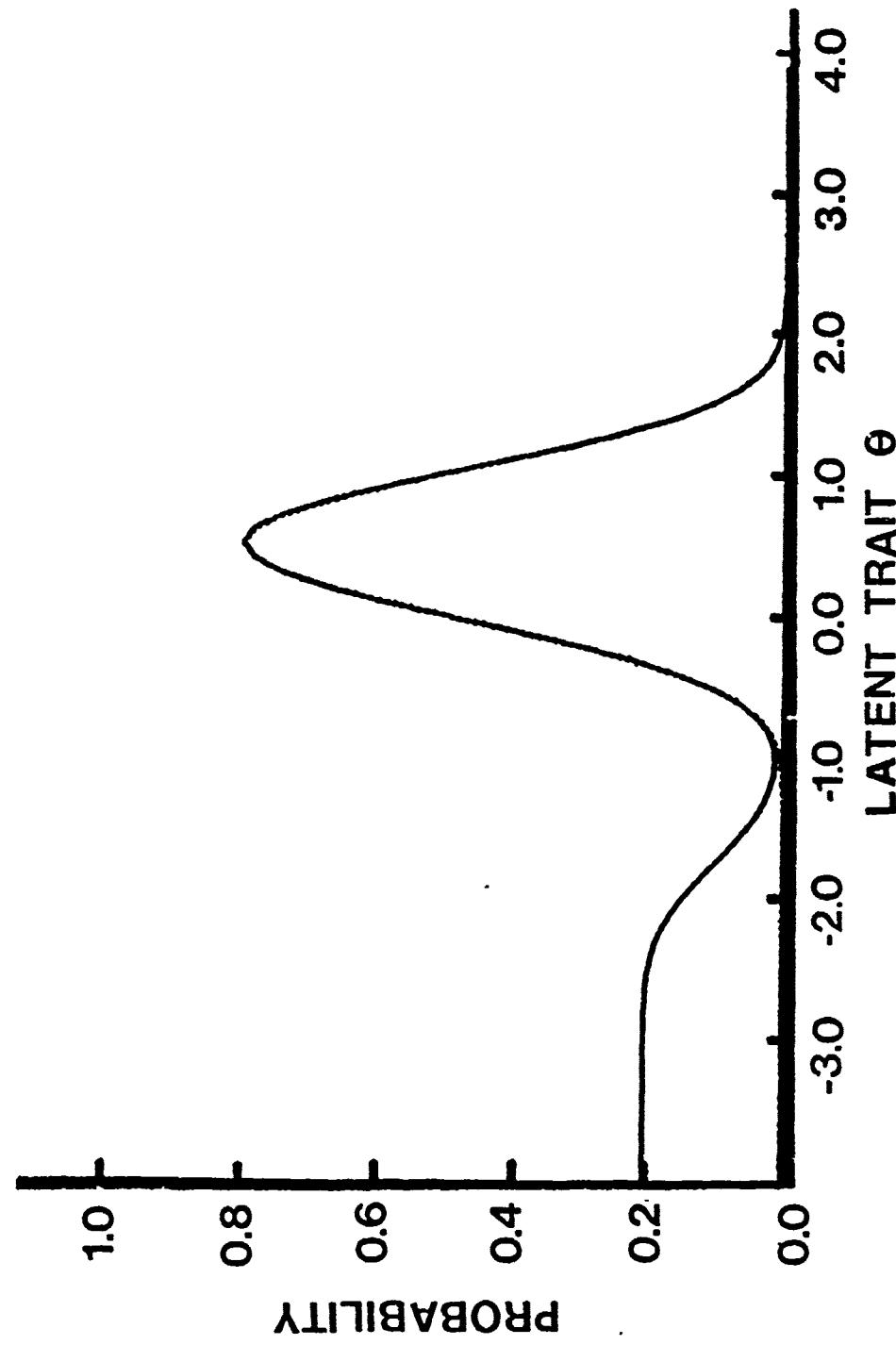


FIGURE A-1-6 (Continued)  $h = 3$ .

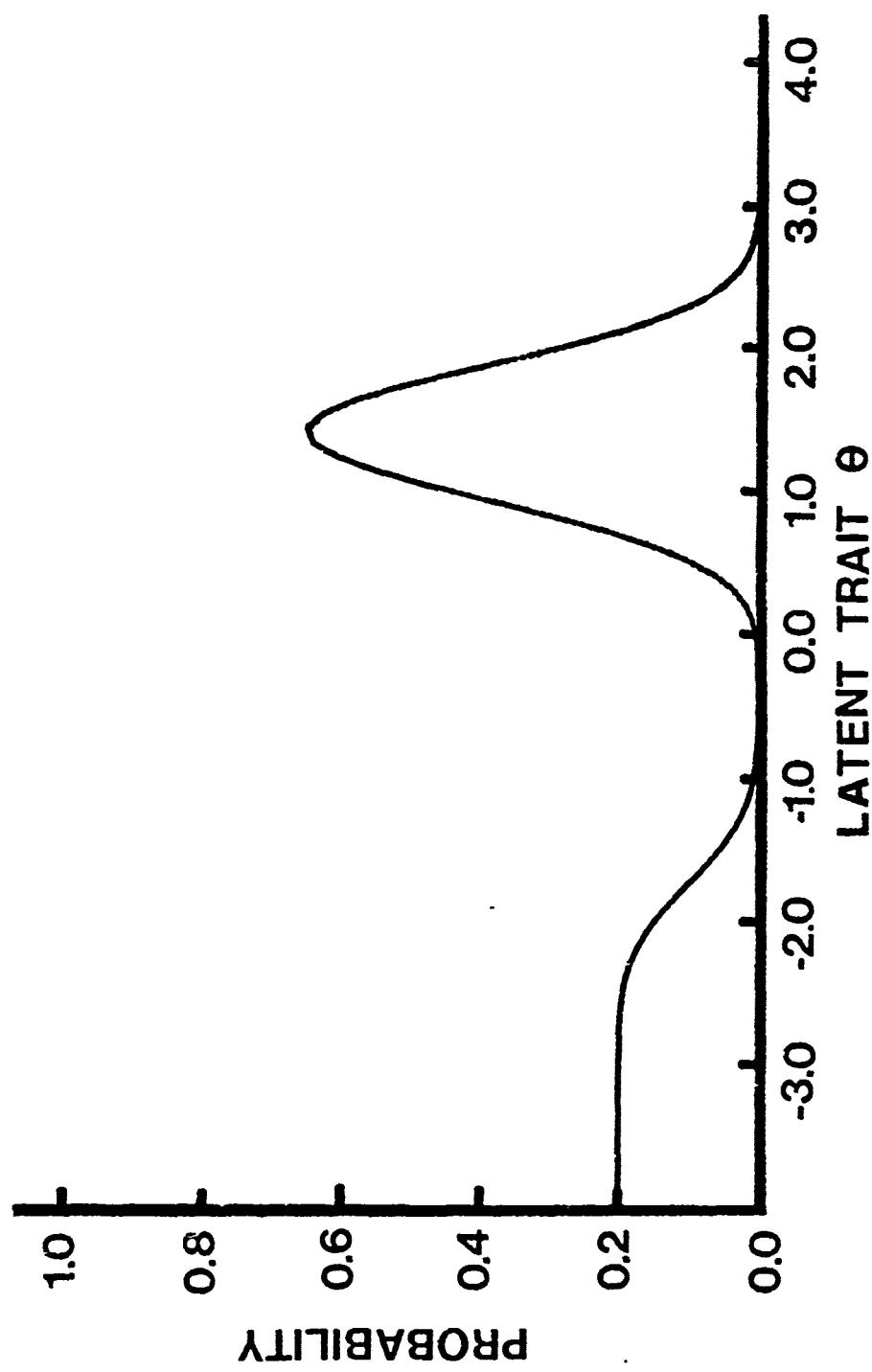


FIGURE A-1-6 (Continued)  $h = 4$ .

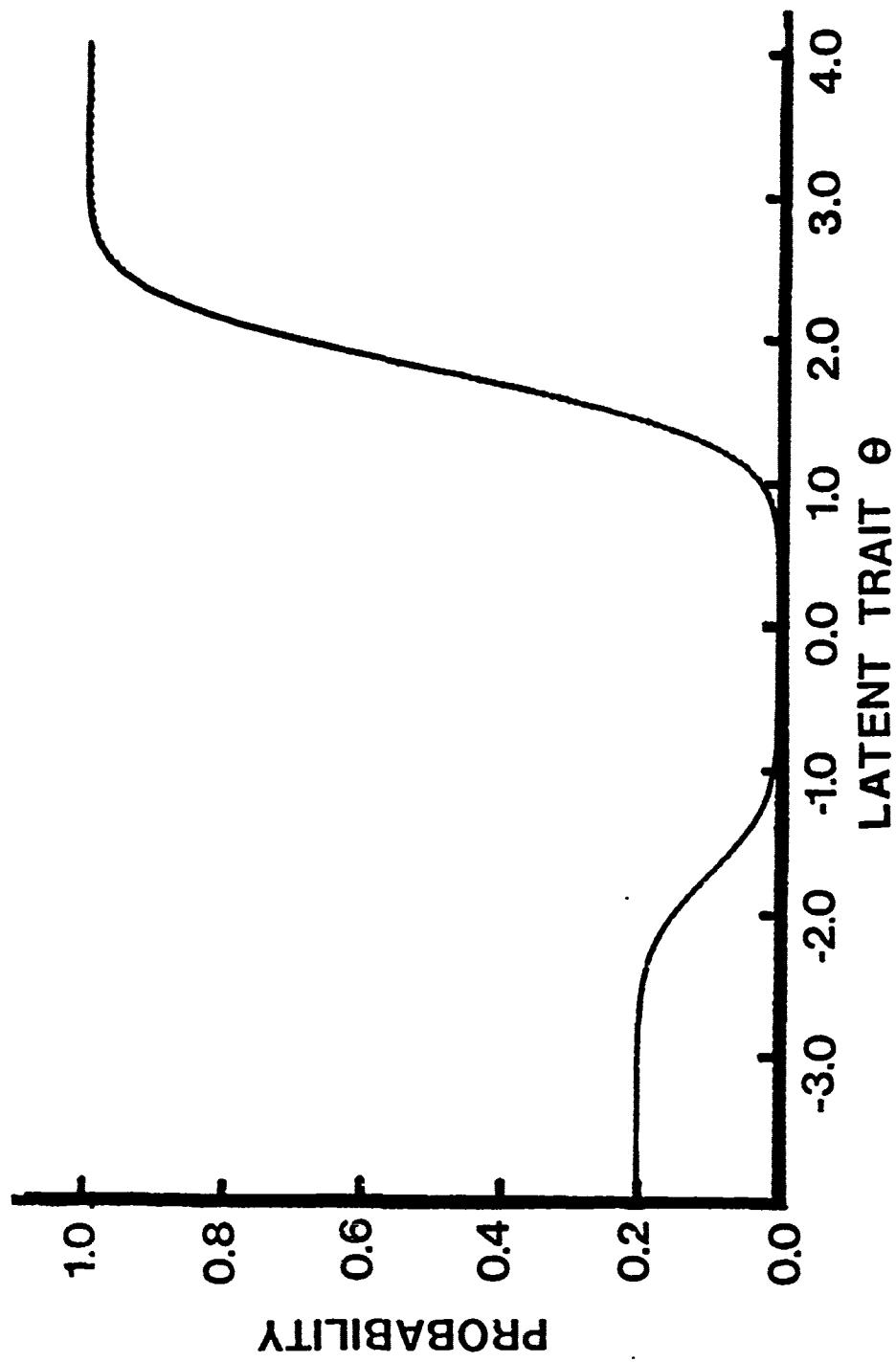


FIGURE A-1-6 (Continued)  $h = 5$ .

APPENDIX II

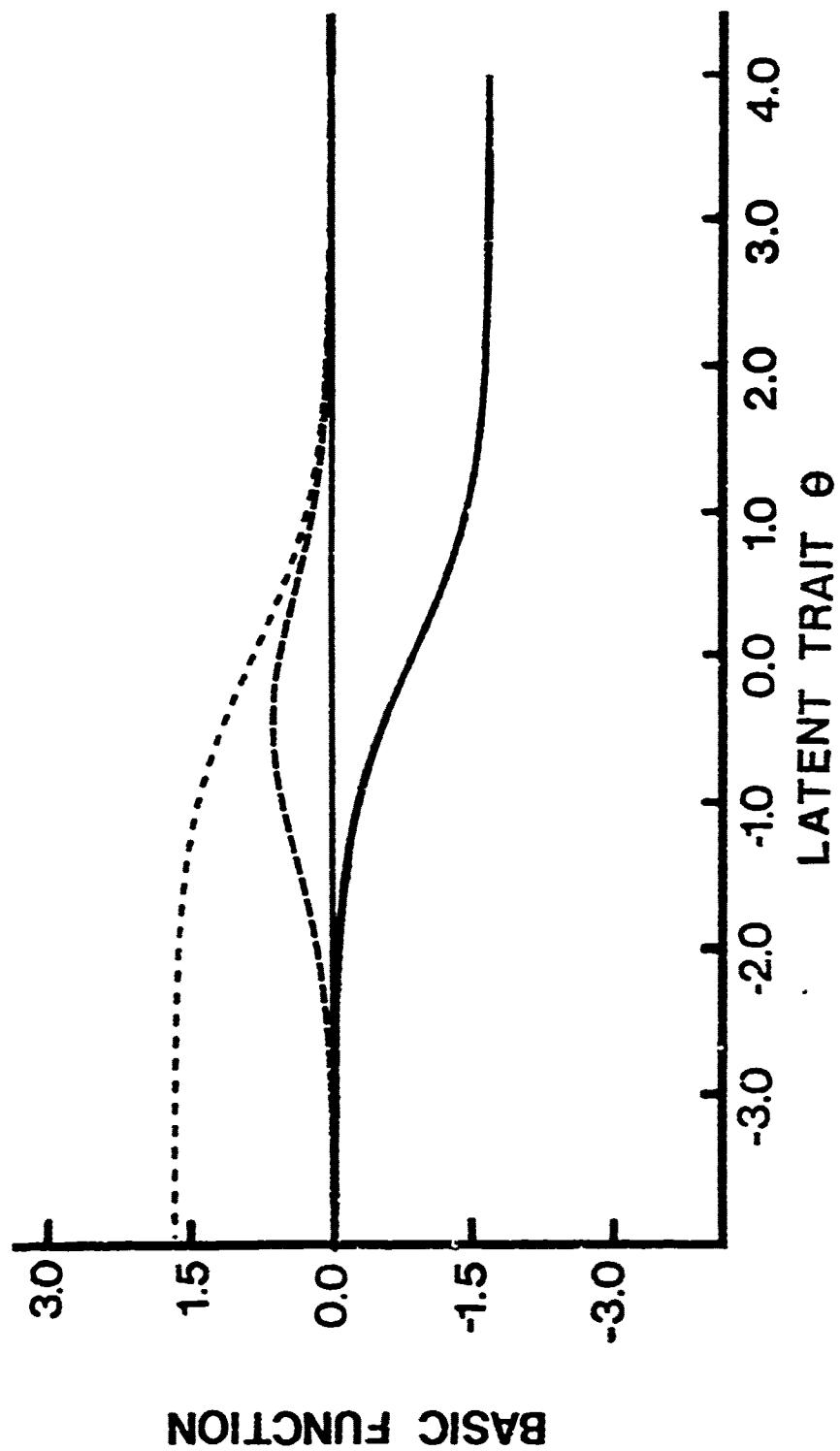


FIGURE A-2-1

Basic Functions of the Two Binary Item Scores Following the Logistic Model (Solid and Dotted Curves) and Those in the Three-Parameter Logistic Model (Solid and Dashed Curves). Those for  $u_g = 0$  (Solid Curves) Are Overlapping Across the Two Models. Item Parameters:  $a_g = 1.00$ ,  $b_g = 0.00$  and  $c_g = 0.20$ .

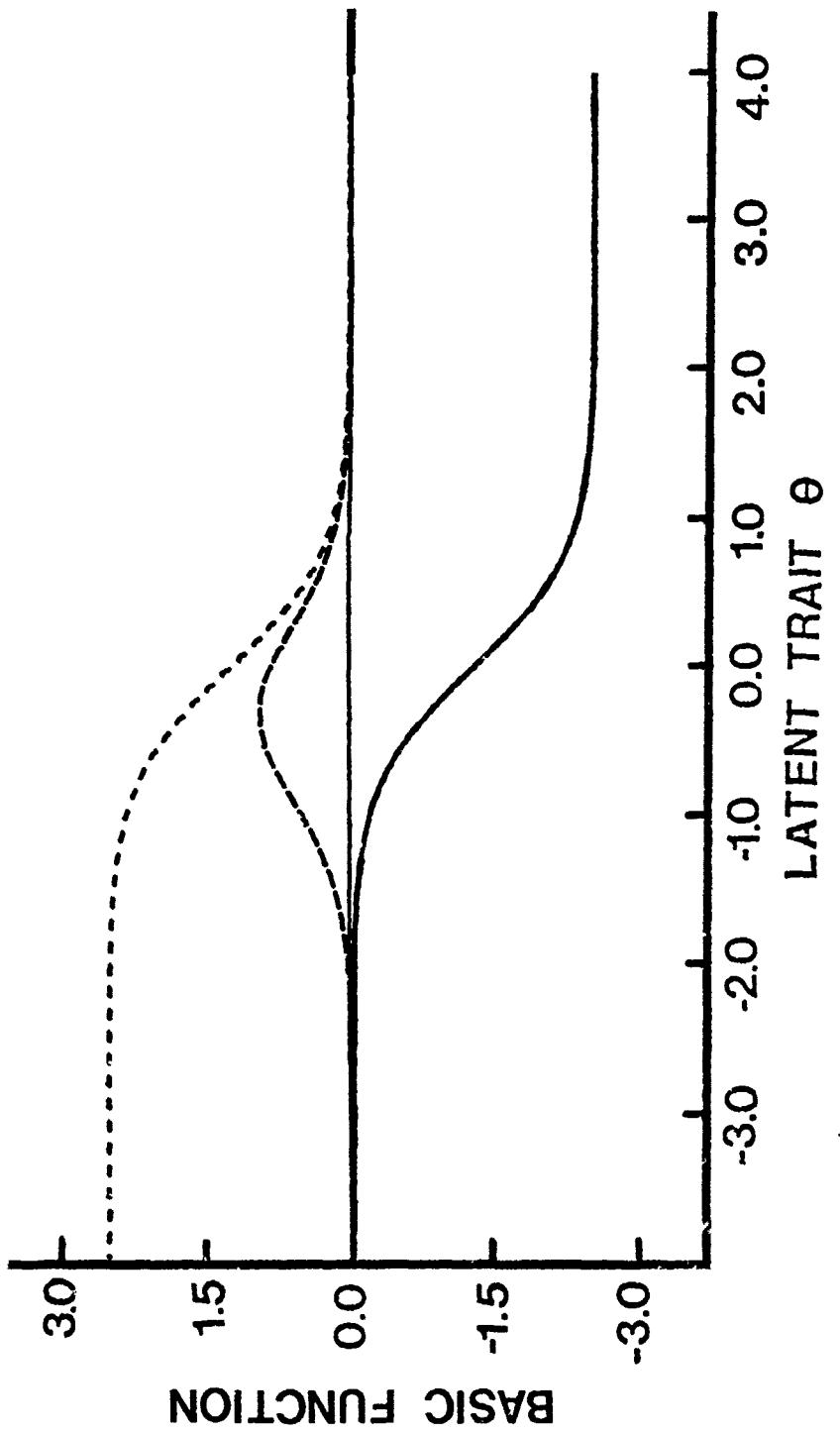


FIGURE A-2-1 (Continued)

Item Parameters:  $a_g = 1.50$ ,  $b_g = 0.00$  and  $c_g = 0.20$ .

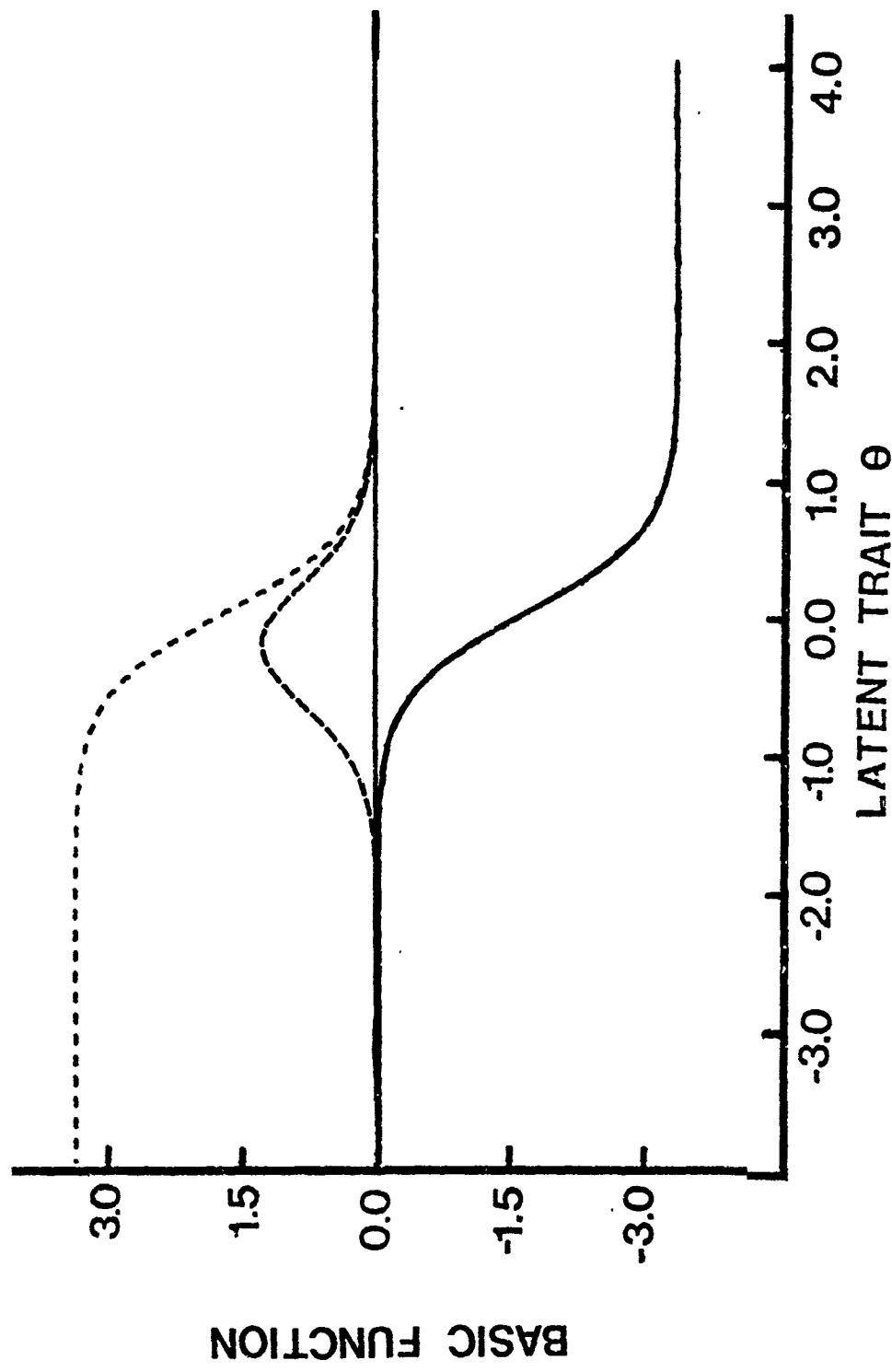


FIGURE A-2-1 (Continued)

Item Parameters:  $a_g = 2.00$ ,  $b_g = 0.00$  and  $c_g = 0.20$ .

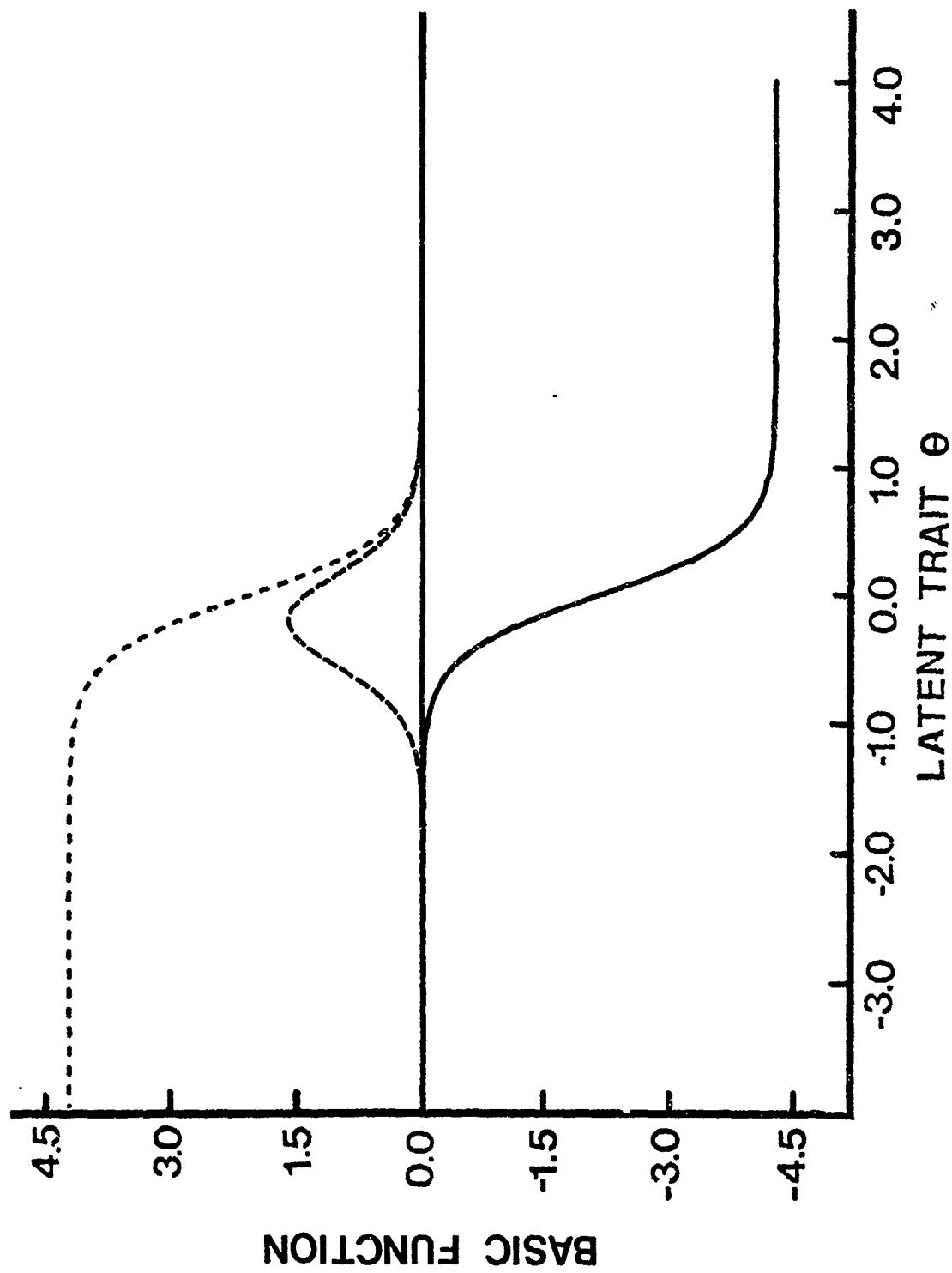


FIGURE A-2-1 (Continued)

Item Parameters:  $a_g = 2.50$ ,  $b_g = 0.00$  and  $c_g = 0.20$ .

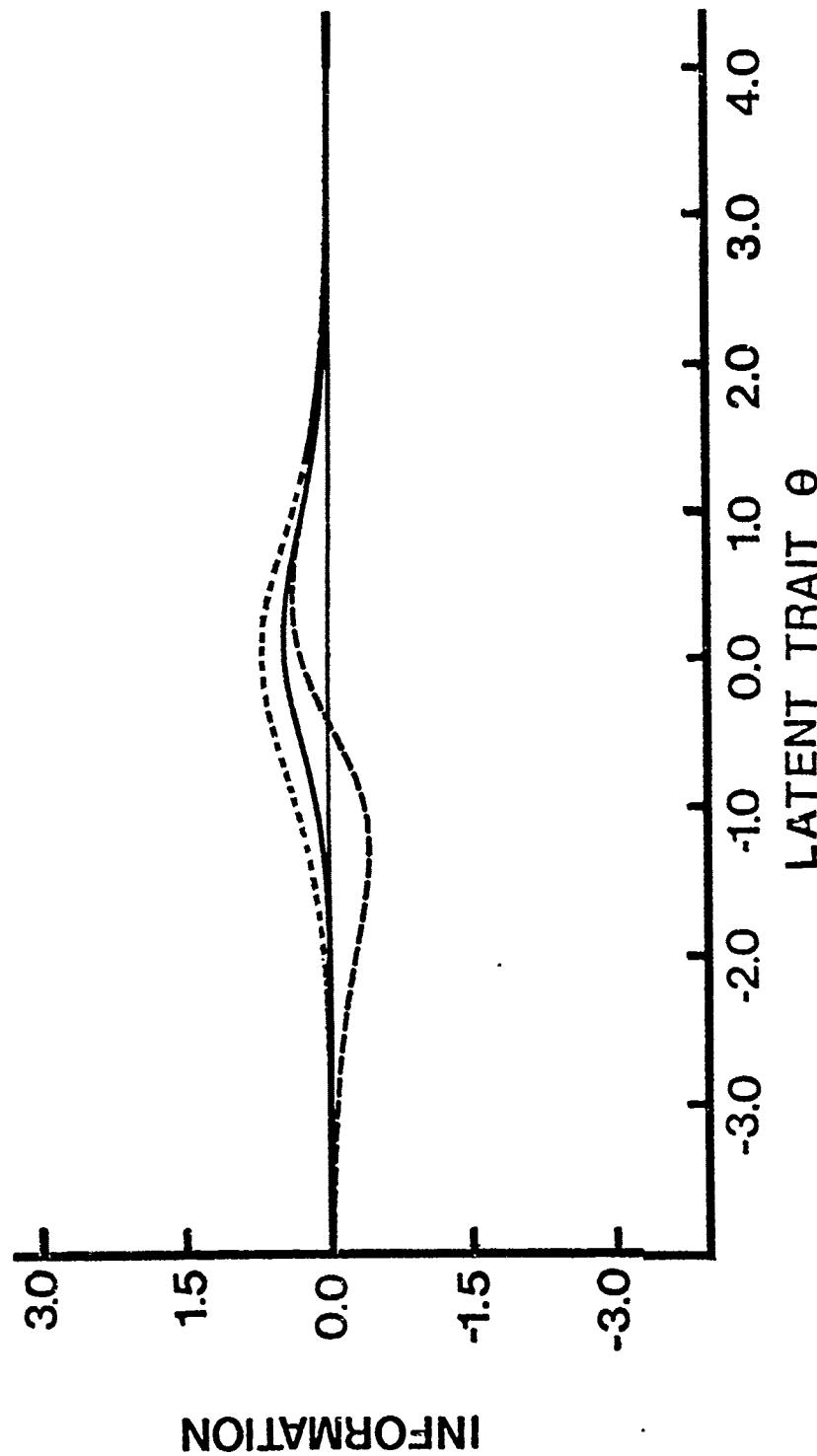


FIGURE A-2-2

Item Response Information Function for  $u_g = 0$  (Dotted Curve) and That for  $u_g = 1$  (Dashed Curve), and Item Information Function (Solid Curve) in the Three-Parameter Logistic Model, and Those in the Logistic Model, Which Are All Identical (Dotted Curves). Item Parameters:  $a_g = 1.00$ ,  $b_g = 0.00$  and  $c_g = 0.20$ .

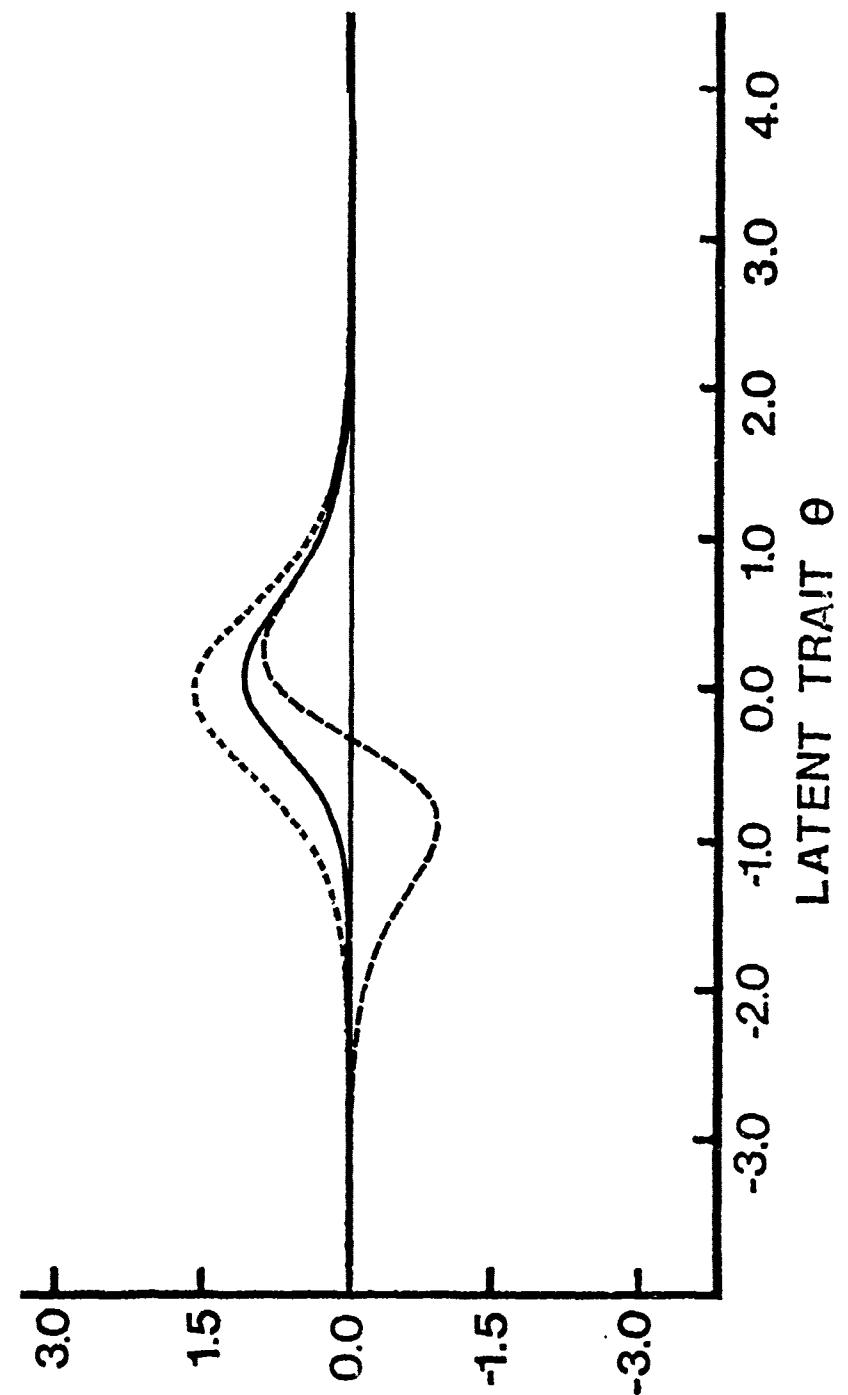


FIGURE A-2-2 (Continued)

Item Parameters:  $a_g = 1.50$ ,  $b_g = 0.00$  and  $c_g = 0.20$ .

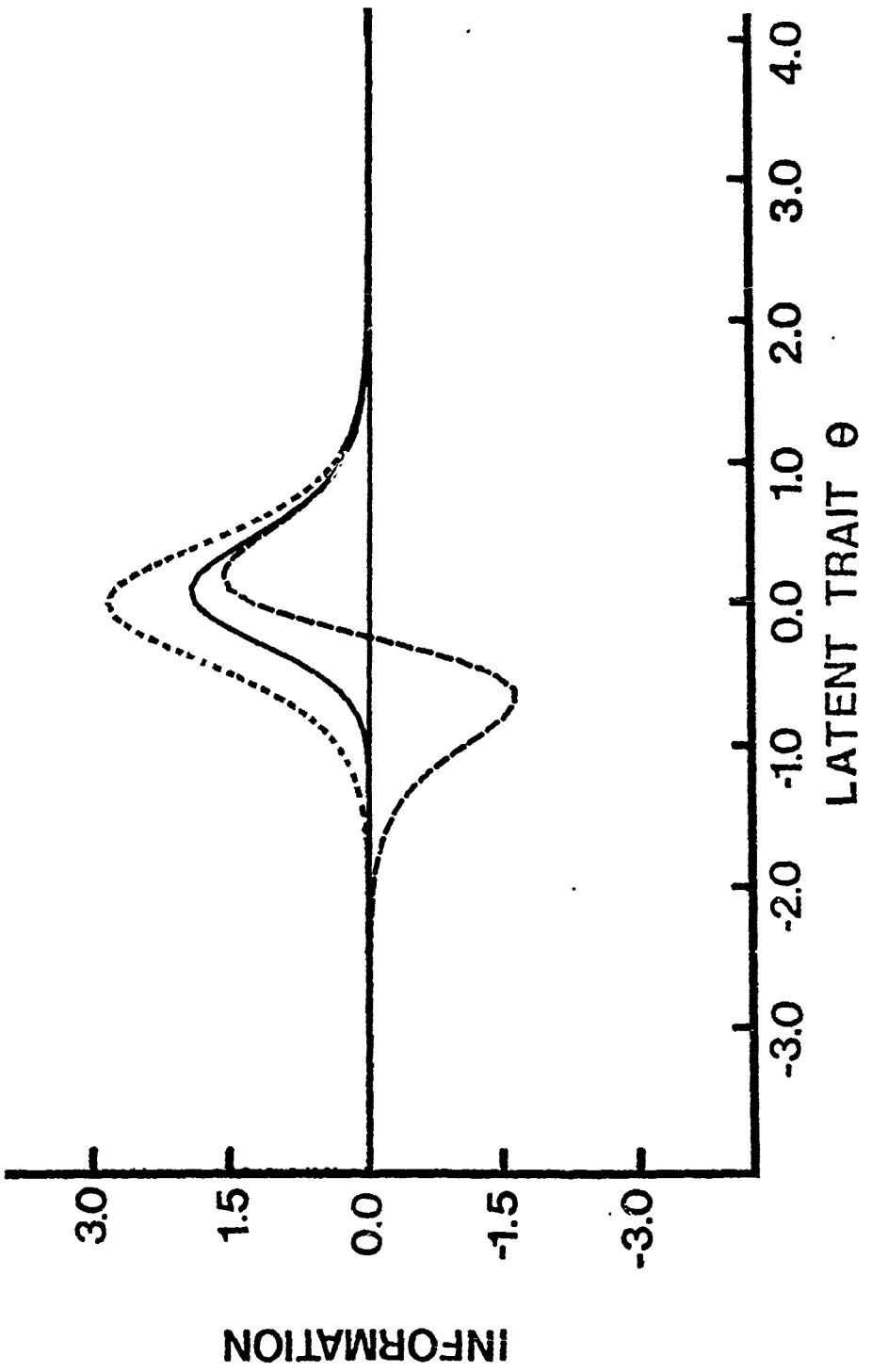


FIGURE A-2-2 (Continued)

Item Parameters:  $a_g = 2.00$ ,  $b_g = 0.00$  and  $c_g = 0.20$ .

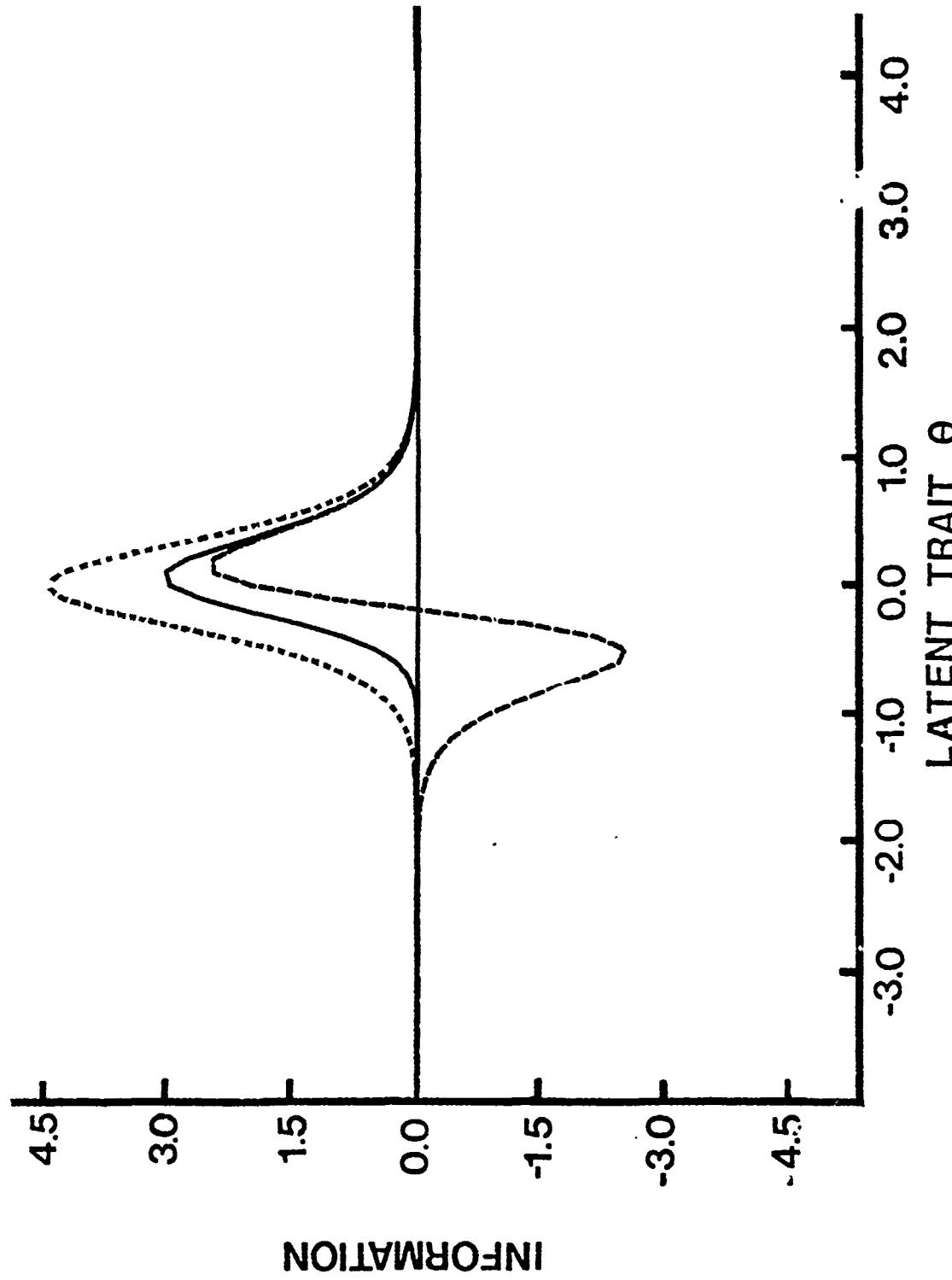


FIGURE A-2-2 (Continued)

Item Parameters:  $a_g = 2.50$ ,  $b_g = 0.00$  and  $c_g = 0.20$ .

DISTRIBUTION LIST

Navy

- 1 Dr. Ed Aiken  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 Dr. Jack R. Borsting  
Provost & Academic Dean  
U.S. Naval Postgraduate School  
Monterey, CA 93940
- 1 Mr. Maurice D. Callahan  
NMPC(5B)  
Navy Military Personnel Command  
Washington, DC 20370
- 1 Dr. Richard Elster  
Department of Administrative Sciences  
Naval Postgraduate School  
Monterey, CA 93940
- 1 DR. PAT FEDERICO  
NAVY PERSONNEL R&D CENTER  
SAN DIEGO, CA 92152
- 1 Dr. Paul Foley  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 Dr. John Ford  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 CAPT. D.M. GRAGG, MC, USN  
HEAD, SECTION ON MEDICAL EDUCATION  
UNIFORMED SERVICES UNIV. OF THE  
HEALTH SCIENCES  
6917 ARLINGTON ROAD  
BETHESDA, MD 20014
- 1 Dr. Norman J. Kerr  
Chief of Naval Technical Training  
Naval Air Station Memphis (75)  
Millington, TN 38054
- 1 Dr. Leonard Kroeker  
Navy Personnel R&D Center  
San Diego, CA 92152

Navy

- 1 CHAIRMAN, LEADERSHIP & LAW DEPT.  
DIV. OF PROFESSIONAL DEVELOPMENT  
U.S. NAVAL ACADEMY  
ANNAPOLIS, MD 21402
- 1 Dr. William L. Maloy  
Principal Civilian Advisor for  
Education and Training  
Naval Training Command, Code OCA  
Pensacola, FL 32508
- 1 Dr. Kneale Marshall  
Scientific Advisor to DCNO(MPT)  
OPO1T  
Washington DC 20370
- 1 CAPT Richard L. Martin  
USS Francis Marion (LPA-249)  
FPO New York, NY 09501
- 1 Dr. James McBride  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 CDR. MERCER  
CNET LIAISON OFFICER  
AFHRL/FLYING TRAINING DIV.  
WILLIAMS AFB, AZ 85224
- 1 Dr. George Moeller  
Head, Human Factors Branch  
Naval Submarine Medical Research Lab  
Groton, CN 06340
- 1 Dr William Montague  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 Commanding Officer  
Naval Health Research  
Center  
Attn: Library  
San Diego, CA 92152
- 1 Naval Medical R&D Command  
Code 44  
National Naval Medical Center  
Bethesda, MD 20014

Navy

- 1 Library  
Navy Personnel R&D Center  
San Diego, CA 92152
- 6 Commanding Officer  
Naval Research Laboratory  
Code 2627  
Washington, DC 20390
- 1 OFFICE OF CIVILIAN PERSONNEL  
(CODE 26)  
DEPT. OF THE NAVY  
WASHINGTON, DC 20390
- 1 DR. JOHN OLSEN  
CHIEF OF NAVAL EDUCATION &  
TRAINING SUPPORT  
PENSACOLA, FL 32509
- 1 Psychologist  
ONR Branch Office  
495 Summer Street  
Boston, MA 02210
- 1 Psychologist  
ONR Branch Office  
536 S. Clark Street  
Chicago, IL 60605
- 1 Office of Naval Research  
Code 200  
Arlington, VA 22217
- 1 Code 436  
Office of Naval Research  
Arlington, VA 22217
- 1 Office of Naval Research  
Code 437  
800 N. Quincy Street  
Arlington, VA 22217
- 5 Personnel & Training Research Programs  
(Code 458)  
Office of Naval Research  
Arlington, VA 22217

Navy

- 1 Psychologist  
OFFICE OF NAVAL RESEARCH BRANCH  
223 OLD MARYLEBONE ROAD  
LONDON, NW, 15TH ENGLAND
- 1 Psychologist  
ONR Branch Office  
1030 East Green Street  
Pasadena, CA 91101
- 1 Scientific Director  
Office of Naval Research  
Scientific Liaison Group/Tokyo  
American Embassy  
APO San Francisco, CA 96503
- 1 Office of the Chief of Naval Operations  
Research, Development, and Studies Branch  
(OP-102)  
Washington, DC 20350
- 1 LT Frank C. Petho, MSC, USNR (Ph.D)  
Code L51  
Naval Aerospace Medical Research Laboratory  
Pensacola, FL 32508
- 1 DR. RICHARD A. POLLAK  
ACADEMIC COMPUTING CENTER  
U.S. NAVAL ACADEMY  
ANNAPOLIS, MD 21402
- 1 Mr. Arnold Rubenstein  
Naval Personnel Support Technology  
Naval Material Command (03T244)  
Room 1044, Crystal Plaza #5  
2221 Jefferson Davis Highway  
Arlington, VA 20360
- 1 DR. A. A. SJOHOLM  
TECH. SUPPORT, CODE 201  
NAVY PERSONNEL R & D CENTER  
SAN DIEGO, CA 92152
- 1 Mr. Robert Smith  
Office of Chief of Naval Operations  
OP-987E  
Washington, DC 20350

1 Navy

1 Dr. Alfred F. Smode  
Training Analysis & Evaluation Group  
(TAEG)  
Dept. of the Navy  
Orlando, FL 32813

1 Dr. Richard Sorensen  
Navy Personnel R&D Center  
San Diego, CA 92152

1 CDR Charles J. Theisen, JR. MSC, USN  
Head Human Factors Engineering Div.  
Naval Air Development Center  
Warminster, PA 18974

1 Dr. W. Gary Thomson  
Naval Ocean Systems Center  
Code 7132  
San Diego, CA 92152

1 Dr. Ronald Weitzman  
Department of Administrative Sciences  
U. S. Naval Postgraduate School  
Monterey, CA 93940

1 DR. MARTIN F. WISKOFF  
NAVY PERSONNEL R& D CENTER  
SAN DIEGO, CA 92152

Army

1 Technical Director  
U. S. Army Research Institute for the  
Behavioral and Social Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 HQ USAREUE & 7th Army  
ODCSOPS  
USAAREUE Director of GED  
APO New York 09403

1 LCOL Gary Bloedorn  
Training Effectiveness Analysis Division  
US Army TRADOC Systems Analysis Activity  
White Sands Missile Range, NM 88002

1 DR. RALPH DUSEK  
U.S. ARMY RESEARCH INSTITUTE  
5001 EISENHOWER AVENUE  
ALEXANDRIA, VA 22333

1 Dr. Myron Fischl  
U.S. Army Research Institute for the  
Social and Behavioral Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 Dr. Ed Johnson  
Army Research Institute  
5001 Eisenhower Blvd.  
Alexandria, VA 22333

1 Dr. Milton S. Katz  
Individual Training & Skill  
Evaluation Technical Area  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 Dr. Beatrice J. Farr  
Army Research Institute (PERI-OK)  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 Dr. Milt Maier  
U.S. ARMY RESEARCH INSTITUTE  
5001 EISENHOWER AVENUE  
ALEXANDRIA, VA 22333

Army

1 Dr. Harold F. O'Neil, Jr.  
Attn: PERI-OK  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 DR. JAMES L. RANEY  
U.S. ARMY RESEARCH INSTITUTE  
5001 EISENHOWER AVENUE  
ALEXANDRIA, VA 22333

1 Dr. Robert Ross  
U.S. Army Research Institute for the  
Social and Behavioral Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 Dr. Robert Sasmor  
U. S. Army Research Institute for the  
Behavioral and Social Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 Director, Training Development  
U.S. Army Administration Center  
ATTN: Dr. Sherrill  
Ft. Benjamin Harrison, IN 46218

1 Dr. Frederick Steinheiser  
U. S. Army Reserch Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

1 Dr. Joseph Ward  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Lab  
AFHRL/PED  
Brooks AFB, TX 78235

1 Air University Library  
AUL/LSE 70/473  
Maxwell AFB, AL 36112

1 Dr. Earl A. Alluisi  
HQ, AFHRL (AFSC)  
Brooks AFB, TX 78235

1 Dr. Philip De Leo  
AFHRL/TT  
Lowry AFB, CO 80230

1 DR. G. A. ECKSTRAND  
AFHRL/AS  
WRIGHT-PATTERSON AFB, OH 45433

1 Dr. Genevieve Haddad  
Program Manager  
Life Sciences Directorate  
AFCSR  
Rolling AFB, DC 20332

1 Dr. Ross L. Morgan (AFHRL/ASR)  
Wright -Patterson AFB  
Ohio 45433

1 Dr. Roger Pennell  
AFHRL/TT  
Lowry AFB, CO 80230

1 Personnel Analysis Division  
HQ USAF/DPXXA  
Washington, DC 20330

1 Research Branch  
AFMPC/DPMYP  
Randolph AFB, TX 78148

1 Dr. Malcolm Ree  
AFHRL/MP  
Brooks AFB, TX 78235

1 Dr. Marty Rockway (AFHRL/TT)  
Lowry AFB  
Colorado 80230

Air Force

- 1 Lt Col Wayne Shore  
Air Force Personnel Center  
Brooks AFB, TX 78235
- 1 Jack A. Thorpe, Maj., USAF  
Naval War College  
Providence, RI 02846
- 1 Dr. Joe Ward, Jr.  
AFHRL/MP  
Brooks AFB, TX 98233
- 1 Brian K. Waters, LCOL, USAF  
Air University  
Maxwell AFB  
Montgomery, AL 36112

Marines

- 1 Dr. H. William Greenup  
Education Advisor (E031)  
Education Center, MCDEC  
Quantico, VA 22134
- 1 Director, Office of Manpower Utilization  
HQ, Marine Corps (MPU)  
BCB, Bldg. 2009  
Quantico, VA 22134
- 1 Major Mike Patro  
Headquarters  
Marine Corps  
Washington, DC 20380
- 1 DR. A.L. SLAFKOSKY  
SCIENTIFIC ADVISOR (CODE RD-1)  
HQ, U.S. MARINE CORPS  
WASHINGTON, DC 20380

CoastGuard

- 1 Mr. Richard Lanterman  
PSYCHOLOGICAL RESEARCH (G-P-1/62)  
J.S. COAST GUARD HQ  
WASHINGTON, DC 20590
- 1 Dr. Thomas Warm  
U. S. Coast Guard Institute  
P. O. Substation 18  
Oklahoma City, OK 73169

Other DoD

- 12 Defense Documentation Center  
Cameron Station, Bldg. 5  
Alexandria, VA 22344  
Attn: TC
- 1 Dr. Craig I. Fields  
Advanced Research Projects Agency  
1400 Wilson Blvd.  
Arlington, VA 22209
- 1 Dr. Dexter Fletcher  
ADVANCED RESEARCH PROJECTS AGENCY  
1400 WILSON BLVD.  
ARLINGTON, VA 22209
- 1 Dr. William Graham  
Testing Directorate  
MEPCOM  
Ft. Sheridan, IL 60037
- 1 Director, Research and Data  
OASD(MRA&L)  
3B919, The Pentagon  
Washington, DC 20301
- 1 Military Assistant for Training and  
Personnel Technology  
Office of the Under Secretary of Defense  
for Research & Engineering  
Room 3D129, The Pentagon  
Washington, DC 20301
- 1 MAJOR Wayne Sellman, USAF  
Office of the Assistant Secretary  
of Defense (MRA&L)  
3B930 The Pentagon  
Washington, DC 20301
- 1 Mr. Fredrick W. Suffa  
NPP (A&R)  
2B269  
Pentagon  
Washington, D.C. 20301

Civil Govt

- 1 Dr. Susan Chipman  
Basic Skills Program  
National Institute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. Lorraine D. Eyde  
Personnel R&D Center  
U.S. Civil Service Commission  
1900 EStreet NW  
Washington, D.C. 20415
- 1 Dr. William Gorham, Director  
Personnel R&D Center  
Office of Personnel Management  
1900 E Street NW  
Washington, DC 20415
- 1 Dr. Joseph I. Lipson  
Division of Science Education  
Room W-638  
National Science Foundation  
Washington, DC 20550
- 1 Dr. John Mays  
National Institute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. Arthur Melmed  
National Intitute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. Andrew R. Molnar  
Science Education Dev.  
and Research  
National Science Foundation  
Jashington, DC 20550
- 1 Dr. Lalitha P. Sanathanan  
Environmental Impact Studies Division  
Argonne National Laboratory  
9700 S. Cass Avenue  
Argonne, IL 60439

Civil Govt

- 1 Dr. Jeffrey Schiller  
National Institute of Education  
1200 19th St. NW  
Washington, DC 20208
- 1 Dr. H. Wallace Sinaiko  
Program Director  
Manpower Research and Advisory Services  
Smithsonian Institution  
801 North Pitt Street  
Alexandria, VA 22314
- 1 Dr. Vern W. Urry  
Personnel R&D Center  
Office of Personnel Management  
1900 E Street NW  
Washington, DC 20415
- 1 DR. C. S. WINIEWICZ  
U.S. CIVIL SERVICE COMMISSION  
REGIONAL PSYCHOLOGIST  
230 S. DEARBORN STREET  
CHICAGO, IL 60604
- 1 Dr. Joseph L. Young, Director  
Memory & Cognitive Processes  
National Science Foundation  
Washington, DC 20550

Non Govt

- 1 Dr. Erling E. Anderson  
University of Copenhagen  
Studiestraedt  
Copenhagen  
DENMARK
- 1 Dr. John Annett  
Department of Psychology  
University of Warwick  
Coventry CV4 7AL  
ENGLAND
- 1 1 psychological research unit  
Dept. of Defense (Army Office)  
Campbell Park Offices  
Canberra ACT 2600, Australia
- 1 Dr. Alan Baddeley  
Medical Research Council  
Applied Psychology Unit  
15 Chaucer Road  
Cambridge CB2 2EF  
ENGLAND
- 1 Dr. Isaac Bejar  
Educational Testing Service  
Princeton, NJ 08450
- 1 Dr. Werner Birice  
Streitkraefteamt  
Rosenberg 5300  
Bonn, West Germany D-5300
- 1 Dr. R. Darrel Bock  
Department of Education  
University of Chicago  
Chicago, IL 60637
- 1 Dr. Nicholas A. Bond  
Dept. of Psychology  
Sacramento State College  
600 Jay Street  
Sacramento, CA 95819
- 1 Dr. David G. Bowers  
Institute for Social Research  
University of Michigan  
Ann Arbor, MI 48106

Non Govt

- 1 Dr. Robert Brennan  
American College Testing Programs  
P. O. Box 168  
Iowa City, IA 52240
- 1 DR. C. VICTOR BUNDERSON  
WICAT INC.  
UNIVERSITY PLAZA, SUITE 10  
1160 SO. STATE ST.  
OREM, UT 84057
- 1 Dr. John B. Carroll  
Psychometric Lab  
Univ. of No. Carolina  
Davie Hall 013A  
Chapel Hill, NC 27514
- 1 Center for the Study of Reading  
174 Children's Research Center  
51 Gerty Drive  
Champiagn, IL 61820
- 1 Charles Myers Library  
Livingstone House  
Livingstone Road  
Stratford  
London E15 2LJ  
ENGLAND
- 1 Dr. Kenneth E. Clark  
Collcge of Arts & Sciences  
University of Rochester  
River Campus Station  
Rochester, NY 14627
- 1 Dr. Norman Cliff  
Dept. of Psychology  
Univ. of So. California  
University Park  
Los Angeles, CA 90007
- 1 Dr. William Coffman  
Iowa Testing Programs  
University of Iowa  
Iowa City, IA 52242

Non Govt

1 Dr. Earl Hunt  
Dept. of Psychology  
University of Washington  
Seattle, WA 98105

1 Dr. Huynh Huynh  
Department of Education  
University of South Carolina  
Columbia, SC 29208

1 Dr. Carl J. Jensema  
Gallaudet College  
Kendall Green  
Washington, DC 20002

1 Dr. John A. Keats  
University of Newcastle  
Newcastle, New South Wales  
AUSTRALIA

1 Mr. Marlin Kroger  
1117 Via Goleta  
Palos Verdes Estates, CA 90274

1 LCOL. C.R.J. LAFLEUR  
PERSONNEL APPLIED RESEARCH  
NATIONAL DEFENSE HQS  
101 COLONEL BY DRIVE  
OTTAWA, CANADA K1A 0K2

1 Dr. Michael Levine  
Department of Educational Psychology  
University of Illinois  
Champaign, IL 61820

1 Dr. Charles Lewis  
Faculteit Sociale Wetenschappen  
Rijksuniversiteit Groningen  
Oude Boteringestraat  
Groningen  
NETHERLANDS

1 Dr. Robert Linn  
College of Education  
University of Illinois  
Urbana, IL 61801

Non Govt

1 Dr. Frederick M. Lord  
Educational Testing Service  
Princeton, NJ 08540

1 Dr. James Lumsdaine  
Department of Psychology  
University of Western Ontario  
Nedlands 6009  
AUSTRALIA

1 Dr. Robert R. Mackie  
Human Factors Research, Inc.  
5775 Dawson Avenue  
Goleta, CA 93017

1 Dr. Gary Marco  
Educational Testing Service  
Princeton, NJ 08450

1 Dr. Scott Maxwell  
Department of Psychology  
University of Houston  
Houston, TX 77025

1 Dr. Sam Mayo  
Loyola University of Chicago  
Chicago, IL 60601

1 Dr. Allen Munro  
Univ. of So. California  
Behavioral Technology Labs  
3717 South Hope Street  
Los Angeles, CA 90007

1 Dr. Melvin R. Novick  
Iowa Testing Programs  
University of Iowa  
Iowa City, IA 52242

1 Dr. Jesse Orlansky  
Institute for Defense Analysis  
400 Army Navy Drive  
Arlington, VA 22202

1 Dr. James A. Paulson  
Portland State University  
P.O. Box 751  
Portland, OR 97207

Non Govt

1 MR. LUIGI PETRULLO  
2431 N. EDGEWOOD STREET  
ARLINGTON, VA 22207

1 DR. STEVEN M. PINE  
4950 Douglas Avenue  
Golden Valley, MN 55416

1 DR. DIANE M. RAMSEY-KLEE  
R-K RESEARCH & SYSTEM DESIGN  
3047 RIDGEMONT DRIVE  
MALIBU, CA 90265

1 MIN. RET. M. RAUCH  
P I I 4  
BUNDESMINISTERIUM DER VERTEIDIGUNG  
POSTFACH 161  
53 BONN 1, GERMANY

1 Dr. Peter B. Read  
Social Science Research Council  
605 Third Avenue  
New York, NY 10016

1 Dr. Mark D. Reckase  
Educational Psychology Dept.  
University of Missouri-Columbia  
12 Hill Hall  
Columbia, MO 65201

1 Dr. Andrew M. Rose  
American Institutes for Research  
1055 Thomas Jefferson St. NW  
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum, Chairman  
Department of Psychology  
Montgomery College  
Rockville, MD 20850

1 Dr. Donald Rubin  
Educational Testing Service  
Princeton, NJ 08540

1 Dr. Larry Ruaner  
Gallaudet College  
Kendall Green  
Washington, DC 20002

Non Govt

1 Dr. J. Ryan  
Department of Education  
University of South Carolina  
Columbia, SC 29208

1 DR. ROBERT J. SEIDEL  
INSTRUCTIONAL TECHNOLOGY GROUP  
HUMRRO  
300 N. WASHINGTON ST.  
ALEXANDRIA, VA 22314

1 Dr. Kazao Shigemasu  
University of Tohoku  
Department of Educational Psychology  
Kawauchi, Sendai 982  
JAPAN

1 Dr. Edwin Shirley  
Department of Psychology  
Florida Technological University  
Orlando, FL 32816

1 Dr. Richard Snow  
School of Education  
Stanford University  
Stanford, CA 94305

1 Dr. Kathryn T. Spoehr  
Department of Psychology  
Brown University  
Providence, RI 02912

1 Dr. Robert Sternberg  
Dept. of Psychology  
Yale University  
Box 11A, Yale Station  
New Haven, CT 06520

1 Dr. Thomas Sticht  
HumRRO  
300 N. Washington Street  
Alexandria, VA 22314

1 DR. PATRICK SUPPES  
INSTITUTE FOR MATHEMATICAL STUDIES IN  
THE SOCIAL SCIENCES  
STANFORD UNIVERSITY  
STANFORD, CA 94305

Non Govt

1 Dr. Hariharan Swaminathan  
Laboratory of Psychometric and  
Evaluation Research  
School of Education  
University of Massachusetts  
Amherst, MA 01003

1 Dr. Brad Sympson  
Office of Data Analysis Research  
Educational Testing Service  
Princeton, NJ 08541

1 Dr. Kikumi Tatsuoka  
Computer Based Education Research  
Laboratory  
252 Engineering Research Laboratory  
University of Illinois  
Urbana, IL 61801

1 Dr. Maurice Tatsuoka  
Department of Educational Psychology  
University of Illinois  
Champaign, IL 61801

1 Dr. David Thissen  
Department of Psychology  
University of Kansas  
Lawrence, KS 66044

1 Dr. Robert Tsutakawa  
Dept. of Statistics  
University of Missouri  
Columbia, MO 65201

1 Dr. J. Uhlaner  
Perceptronics, Inc.  
6271 Variel Avenue  
Woodland Hills, CA 91364

1 Dr. Howard Wainer  
Bureau of Social SCience Research  
1990 M Street, N. W.  
Washington, DC 20036

1 Dr. John Wannous  
Department of Management  
Michigan University  
East Lansing, MI 48824

Non Govt

1 Dr. David J. Weiss  
N660 Elliott Hall  
University of Minnesota  
75 E. River Road  
Minneapolis, MN 55455

1 DR. SUSAN E. WHITELY  
PSYCHOLOGY DEPARTMENT  
UNIVERSITY OF KANSAS  
LAWRENCE, KANSAS 66044

1 Dr. William B. Whitten, II  
Department of Psychology  
SUNY, Albany  
1400 Washington Avenue  
Albany, NY 12222

1 Dr. Wolfgang Wildgrube  
Streitkraefteamt  
Rosenberg 5300  
Bonn, West Germany D-5300

1 Dr. J. Arthur Woodward  
Department of Psychology  
University of California  
Los Angeles, CA 90024

1 Dr. Robert Woud  
School Examination Department  
University of London  
66-72 Gower Street  
London WCIE 6EE  
ENGLAND

Navy

1 Dr. Donald Calder  
Office of Naval Research  
325 Hinman Research Building  
Atlanta, GA 30332

Army

1 Dr. Randall M. Chambers, Ph.D.  
U.S. Army Research Institute  
for the Behavioral and Social  
Sciences  
Fort Sill Field Unit  
P. O. Box 3066  
Fort Sill, OK 73503

Non Govt

1 Dr. Ron Hambleton  
School of Education  
University of Massachusetts  
Amherst, Mass. 01002

1 Dr. William E. Coffman  
Iowa Testing Programs  
334 Lindquist Center  
Iowa City, IA 52242

Non Govt

1 Mr. Isaac I. Bejar  
Department of Psychology  
Elliott Hall  
75 East River Road  
Minneapolis, Minnesota 55455

1 Mr. George Woods  
1106 Newport Ave.  
Victoria, B. C.  
V8S 5E4 Canada

1 Dr. P. Mengal  
Faculte' de Psychologie  
et des Sciences de l'Education  
Universite' de Geneve  
3 fl. de l'Universite  
1201 Geneva SWITZERLAND

1 Dr. Wim J. van der Linden  
Vakgroep Onderwijskunde  
Postbus 217  
7500 EA Enschede  
The Netherlands

1 Dr. Lowell Schipper  
Department of Psychology  
Bowling Green State University  
Bowling Green, Ohio 43403

Dr. Gerhard Fischer  
Liebigasse 5  
Vienna 1010, AUSTRIA

Dr. Lutz Hornke  
University Duesseldorf  
Erz. Wiss.  
D-4000 Duesseldorf  
WEST GERMANY

Dr. Wolfgang Buchtala  
8346 Simbach Inn  
Postfach 1306  
Industriestrasse 1  
WEST GERMANY